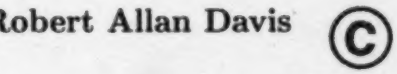
University of Alberta

A Search for a Neutral Higgs Boson in Electron Positron Collisions at LEP

by

Robert Allan Davis 

A thesis submitted to the Faculty of Graduate Studies and Research in partial fulfillment of the requirements for the degree of Doctor of Philosophy

Department of Physics

Edmonton, Alberta

Fall 1999

National Library of Canada
Acquisitions and Bibliographic Services
395 Wellington Street
Ottawa ON K1A 0N4
Canada

Bibliothèque nationale du Canada
Acquisitions et services bibliographiques
395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file Votre référence

Our file Notre référence

The author has granted a non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of this thesis in microform, paper or electronic formats.

The author retains ownership of the copyright in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

L'auteur a accordé une licence non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de cette thèse sous la forme de microfiche/film, de reproduction sur papier ou sur format électronique.

L'auteur conserve la propriété du droit d'auteur qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

0-612-46827-5

Canada

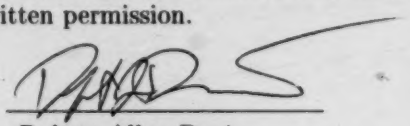
University of Alberta

Library Release Form

NAME OF AUTHOR:	Robert Allan Davis
TITLE OF THESIS:	A Search for a Neutral Higgs Boson in Electron Positron Collisions at LEP
DEGREE:	Doctor of Philosophy
YEAR THIS DEGREE GRANTED:	1999

Permission is hereby granted to the University of Alberta Library to reproduce single copies of this thesis and to lend or sell such copies for private, scholarly or scientific research purposes only.

The author reserves all other publication and other rights in association with the copyright in the thesis, and except as hereinbefore provided neither the thesis nor any substantial portion thereof may be printed or otherwise reproduced in any material form whatever without the author's prior written permission.



Robert Allan Davis
9825 - 92 Avenue NW,
Edmonton, Alberta
T6E 2V4

October 4, 1999

University of Alberta

Faculty of Graduate Studies and Research

The undersigned certify that they have read, and recommend to the Faculty of Graduate Studies and Research for acceptance, a thesis entitled **A Search for a Neutral Higgs Boson in Electron Positron Collisions at LEP** submitted by **Robert Allan Davis** in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

Professor James L. Pinfold
(Supervisor)

Professor Claude Leroy
Université de Montréal
(External)

Professor Helmy S. Sherif
(Chairman)

Assistant Professor Frank A. Hegmann

Assistant Professor Marc de Montigny
Faculté Saint-Jean

Assistant Professor Manuella G. Vincter

DATE: October 1, 1999

This thesis is dedicated to Julie and Marc, with all my love forever.

Abstract

A search for a Standard Model Higgs boson in the $e^+e^- \to h^0_{SM}Z^0 \to b\bar{b}q\bar{q}$ channel is presented using data collected at a centre of mass energy of 183 GeV by the OPAL experiment in 1997. The analysis is also extended to search for the lightest neutral scalar Higgs boson of the Minimal Supersymmetric Model (MSSM). The search algorithm makes use of an Artificial Neural Network for data classification. For the Standard Model analysis, seven events were selected in the data which is consistent with the $6.15 \pm 0.52(\text{stat.}) \pm 1.16(\text{syst.})$ events expected from background processes alone. If a Standard Model Higgs boson with a mass of 85 GeV/c^2 were to exist, the analysis would be sensitive to $6.44 \pm 0.04(\text{stat.}) \pm 0.49(\text{syst.})$ events. Combining the results of the Standard Model analysis with the other decay channels studied by OPAL yields a lower mass limit of 86.94 GeV/c^2 at the 95% confidence level for a Standard Model Higgs boson. Combining the supersymmetric extension of the analysis with the other search channels studied in OPAL gives a lower mass limit of 67.5 GeV/c^2 for both the lightest neutral scalar Higgs boson and the CP-odd neutral Higgs boson in a constrained Minimal Supersymmetric Model.

Preface

This thesis documents research performed at the University of Alberta between May 1997 and August 1999.

The Standard Model results from chapter 4 represent the latest evolution of an original application of Artificial Neural Networks (ANN) to high energy physics analysis. Earlier ANN analyses by the author were the first application of this technique in the OPAL Higgs group at CERN. Since then, at least two other analyses make use of ANN's in the Higgs group. The results of chapter 4 have been published in the following article:

OPAL Collab., G. Abbiendi *et al.*, *Search for Higgs Bosons in* e^+e^- *Collisions at 183 GeV*, Euro. Jour. Physics **C7** (1999) 407-435.

During the past two and one half years, the author also contributed to the construction and testing of the Time of Flight Endcap (TE) detector in OPAL. The results of this work are documented in the following article:

G. Aguillion *et al.*, *Thin Scintillating Tiles with High Light Yield for the OPAL Endcaps*, Nucl. Instrum. Methods **A417** (1998) 266-277.

Finally, the author was responsible for the installation and maintenance of all OPAL related software and computing facilities at the University of Alberta during his tenure as both a MSc. student and a Ph.D. student.

Acknowledgements

I have many people to thank, probably more than are listed below, for their help and support during my time at the University of Alberta. In particular, I would like to thank:

- Dr. James L. Pinfold for his support of the research in this thesis and for providing a challenging environment where I could develop my skills.
- Dr. Doug Gingrich and Dr. Nate Rodning for many useful and enjoyable discussions, both on and off the topic of physics.
- Audrey Schaapman for all of her help with the day to day running of the Centre for Subatomic Research.
- Jim MacKinnon and Jay Haverstock for keeping the system rolling.
- My family for all of their support through my many years in school.
- And finally to Julie, who has taught me so much.

Contents

1 Introduction and Theory 1
- 1.1 A Survey of the Fundamental Particles and Forces . . . 2
  - 1.1.1 The Forces . . . 2
  - 1.1.2 The Particles . . . 3
  - 1.1.3 The Mathematical Framework . . . 4
- 1.2 The Higgs Mechanism in the Standard Model . . . 6
  - 1.2.1 The Gauge Boson Masses . . . 8
  - 1.2.2 The Fermion Masses . . . 9
  - 1.2.3 Properties of the Higgs Boson . . . 11
  - 1.2.4 Production of Higgs Bosons at LEP . . . 12
- 1.3 Successes and Deficiencies of the Standard Model . . . 15
- 1.4 The Minimal Supersymmetric Model . . . 16
  - 1.4.1 Motivation: The Gauge Hierarchy Problem . . . 17
  - 1.4.2 The Higgs Mechanism in the MSSM . . . 18
  - 1.4.3 The MSSM Scan at OPAL . . . 22

2 The Experimental Apparatus 25
- 2.1 The Large Electron Positron Collider . . . 26
- 2.2 The OPAL Detector . . . 30
  - 2.2.1 The Beam Pipe and Magnet . . . 31
  - 2.2.2 The Central Detector . . . 33
  - 2.2.3 The Time of Flight (TOF) System . . . 45
  - 2.2.4 The Calorimeter System . . . 48
  - 2.2.5 The Muon System . . . 53
  - 2.2.6 The Forward Detector System . . . 54
  - 2.2.7 Triggering and Online Data Acquisition . . . 55

3 Artificial Neural Networks 58
- 3.1 The Neuron . . . 59
- 3.2 Structure of an ANN . . . 61
- 3.3 Significance of Hidden Neurons . . . 63
- 3.4 The Learning Process . . . 66
- 3.5 Backpropagation . . . 68

3.6 Considerations for Successful ANN's 69
3.7 Applications of ANN's in Particle Physics 72
3.7.1 The CDF Isolation Trigger 72
3.7.2 Vertex Finding with Impact Parameters 74

4 The Search for a Neutral Higgs Boson **77**
4.1 Search Strategy 78
4.2 Production of Simulated Data 79
4.3 The 183 GeV Data from OPAL 81
4.4 Offline Data Analysis 82
4.4.1 Data Quality 82
4.4.2 Jet Finding 85
4.4.3 Energy Flow Correction 86
4.4.4 Event Shape Variables 87
4.4.5 Effective Centre of Mass Energy 88
4.4.6 B Jet Tagging 89
4.5 Precut 94
4.6 The Higgs Search ANN 97
4.6.1 Input Variables 97
4.6.2 ANN Training 100
4.7 Results of the Standard Model Search 102
4.7.1 Evaluation of Systematic Errors 106
4.7.2 Comparison with Other Analysis Methods 110
4.7.3 Mass Limits for the SM Higgs Boson 110
4.8 Results of the MSSM Search 114

5 Summary and Conclusion **115**

Bibliography **118**

A The OPAL Collaboration **122**

List of Tables

1.1 The fundamental forces and mediating gauge bosons of the Standard Model 2
1.2 Properties of the Standard Model Fermions 4
1.3 Coupling constants of the MSSM Higgs bosons to fermions. 21

2.1 Summary of the OPAL subdetectors. 32

4.1 Monte Carlo data sets used in the search for the Higgs boson. 81
4.2 Summary of OPAL data at 183 GeV 82
4.3 Detector status bits in OPAL. 85
4.4 Detector status bits required for real data in this analysis. 85
4.5 Effectiveness of the precuts. 96
4.6 Final background expectation for the ANN analysis. 105
4.7 Efficiencies for various Higgs masses for the ANN analysis 105
4.8 Candidate events for the ANN four jet analysis. The ANN output and the reconstructed Higgs mass for each candidate is shown. 106

List of Figures

1.1 The Higgs potential. The minimum of the potential (green circle) has no preferred direction. 7
1.2 Branching ratios for Standard Model Higgs boson decay to fermions. . 12
1.3 Branching ratios for Standard Model Higgs boson decay to bosons. . 13
1.4 The Higgstrahlung production mechanism. 13
1.5 Cross section for Higgstrahlung production versus Higgs mass for various centre of mass energies. 14
1.6 Estimate of the Higgs mass from fits to electroweak data. The χ^2 of the fit favours a Higgs mass of about 100 GeV/c^2. 18
1.7 The MSSM exclusion prior to the 183 GeV analysis 24

2.1 The CERN accelerator complex near Geneva, Switzerland. 27
2.2 The four LEP experiments at the their respective interaction points. The OPAL detector is located at Interaction Point 6. 28
2.3 Integrated luminosity delivered to OPAL from 1991 to 1997. 29
2.4 A cutaway view of the OPAL detector. 31
2.5 Cutaway view of the phase three silicon microvertex detector. 35
2.6 Cross-sectional view of the silicon microvertex detector showing the overlap in the detector wafers. 36
2.7 Schematic diagram of a "long" three wafer silicon microvertex detector ladder. 37
2.8 Impact parameter resolution for dimuon events collected at the Z peak. The plot shows the resolution with only the jet chamber (top), the jet chamber plus vertex chamber (middle), and all three central detectors in concert (bottom). 39
2.9 Schematic drawing of the OPAL central vertex chamber. 40
2.10 Wiring layout of the central vertex detector. The page is in the $r\varphi$ plane. 41
2.11 Efficiency for double hit detection of closely spaced tracks in the jet chamber. The efficiency is 80% for track spacings above 2.5 mm for hits alone (dashed line) and 80% for spacings above 3.5 mm for hits associated with a track (solid line). 43

2.12 The momentum resolution of dimuon events in OPAL versus year of running for the jet chamber alone (CJ), jet chamber plus vertex chamber (CJ+CV), and jet chamber, vertex chamber, and silicon microvertex detector in concert (CJ+CV+SI). 44
2.13 Energy loss per unit distance (dE/dx) for various particle types in the OPAL jet chamber. The expected behaviour (lines) and actual measurements (points) are shown for 1994 OPAL data. 46
2.14 A cross-sectional view of a single z chamber. 47
2.15 The Time of Flight Endcap detector. The position of the subdetector in OPAL is shown as well as a schematic of a single detector sector. . 49
2.16 The tile layout of the minimum ionizing particle (MIP) plug. 56

3.1 The dot product neuron. The neuron calculates the dot product of the states of the incoming neurons s_i with their respective weights w_{ij} and then uses this value to calculate the output of the neuron through the activation function f. 60
3.2 The step activation function (left) and the nonlinear sigmoid function (right). 61
3.3 The architecture of a feed forward network showing the input, hidden, and output layers. 62
3.4 The architecture of a recurrent network. Every neuron has a connection to every other neuron in the network. 63
3.5 A simple example of a two layer ANN used to separate two classes of data $\mathcal{A}$ and $\mathcal{B}$ (left). With no hidden layer, the ANN is capable of only a one dimensional cut in feature space along the line $aX + bY + c = 0$. The network architecture is shown on the right. The threshold input node always reports a value of c to the network. 64
3.6 The two dimensional space containing the outputs of A.XOR.B is shown at the top of the diagram. Points for which A.XOR.B is true are shown as dark circles, while points for which A.XOR.B is false are shown as open circles. In the top left and top middle plots, it is shown that a one-dimensional decision surface is insufficient to separate the data classes. The addition of a hidden layer, however, that determines the state of A and B together, makes classification possible. The chart at the bottom of the figure shows the neuron outputs for various inputs A and B. 65
3.7 A simple two layer ANN. 67
3.8 A three layer ANN with inter-neuron weights $w(i,j) \equiv w_{ij}$ 70
3.9 Selection of calorimeter clusters for the CDF isolation trigger ANN. The energy in each of the "inner" (shaded) areas is compared to the energy in the outer area of the 5x5 cell cluster in order to identify isolated leptons in the second level trigger. 73

3.10 Geometric diagram showing how tracks originating from a common secondary vertex give rise to straight lines with the equation $D = R \cdot sin(\arctan \frac{r_x}{z_v} - \Phi)$ in D/Φ space. 74
3.11 Impact parameter versus φ plots for events containing secondary vertices (right) and minimum bias background events (left) at CDF. Note the existence of identifiable straight lines in the left plot. 75
3.12 Identification efficiency of b vertices versus the cut made on the neural network output in CDF. 76

4.1 Cross sections for various background processes to the Higgs signal at LEP versus centre of mass energy. 80
4.2 Merging efficiency for jet chamber and silicon microvertex tracks before and after the CV incident using the original and improved matching algorithms. With the improved matching algorithm, merging efficiency is virtually unaffected after the incident. 83
4.3 Schematic diagram of the procedure used to calculate the jet-wise b-tagging variable $\mathcal{B}_j$. The constituent parts are explained in the text. 90
4.4 Precut distributions for signal and two fermion background. 95
4.5 Precut distributions for signal and four fermion background. 96
4.6 Distributions of the ANN input variables 99
4.7 Mean squared error for the training (top left) and test (top right) sets versus epoch for the ANN training. The bottom two plots show the mean saturation of the nodes of the hidden layer (left) and the mean saturation of the nodes of the output layer (right). The developing plateau in both of these plots indicates that the network has learned its full capacity of information. 101
4.8 Efficiency versus training epoch for 80 GeV/c^2 Higgs events in the test set. The event was selected as signal if the ANN output was greater than 0.7. 103
4.9 Output of the three ANN output nodes (columns) for the signal (first row), Z/γ^* background (second row), and four fermion background (third row). 104
4.10 Output of the first ANN node for the background channels, the total background, and the total background plus an amount of simulated 85 GeV Higgs signal normalized to the integrated luminosity. The ANN output of OPAL data is also shown by the points. The cut on the ANN output at 0.85 is shown and the enhancement of the signal above the cut is visible. 107
4.11 A candidate event selected by the Artificial Neural Network selection used in this thesis. The ANN output for this event was 0.9 and the reconstructed Higgs mass was 82.9 GeV/c^2. 108

4.12 Relation between the mass distribution $D(m)$, the filter function $F(m)$, and the probability $P_1(w)\Delta w$ (hatched area) for one candidate event to have a weight $w \pm \Delta w/2$. In the example chosen $w \pm \Delta w/2 = 0.7 \pm 0.1$. 112
4.13 The MSSM exclusion for the 183 GeV analysis 114

Chapter 1

Introduction and Theory

To date, all available experimental data agrees with the predictions of the Standard Model [1]. With the recent discovery of the top quark at Fermilab near Chicago, only two of the basic ingredients of the Standard Model have yet to be experimentally detected. The search for one of these particles, the Higgs Boson, is the subject of this thesis. As will be seen in the sections that follow, the Higgs Boson is of crucial importance as it exists as a result of the mechanism used to introduce particle masses into the Standard Model.

Despite its success, several questions are left unanswered by the Standard Model, and it is generally believed that it is only a subset of a more fundamental theory. The rest of this chapter presents the ingredients of the Standard Model as well as some of the questions left unanswered by the theory. In particular, the Higgs Mechanism, which is necessary to allow masses to be introduced into the theory, is discussed in some detail. One of the questions left unanswered by the Standard Model serves as a motivation for a possible extension to the theory, the Minimal Supersymmetric Model (MSSM), which is described in section 1.4.

Force	Relative Strength	Range (m)	Gauge Boson	Boson Mass ($/c^2$)	Boson Spin	Acts On
Strong Force	≈ 1	$< 10^{-15}$	Gluons (8)	0	1	Coloured Particles
Electro-magnetism	7.3×10^{-3}	∞	Photon (γ)	0	1	Charged Particles
Weak Force	1.02×10^{-5}	10^{-18}	$W^{\pm}$ Z^0	80 GeV 91 GeV	1 1	Quarks, Leptons, Gauge Bosons
Gravity	0.53×10^{-38}	∞	Graviton	0	2	All Particles

Table 1.1: The fundamental forces and mediating gauge bosons of the Standard Model

1.1 A Survey of the Fundamental Particles and Forces

The Standard Model exhibits both radical change from and remarkable continuity with the field of particle physics prior to 1970. The hundreds of seemingly unrelated particles discovered since the middle of the century have been organized into a consistent theory comprised of a small number of fundamental particles and forces. Quarks and leptons, which belong to the class of spin 1/2 particles called fermions, comprise the fundamental particles of which matter is made. These particles interact with each other through the exchange of integral spin gauge bosons that mediate the four fundamental forces of the theory.

In the sections that follow, the forces and particles of the Standard Model are presented as is the underlying mathematics that binds the theory together.

1.1.1 The Forces

The four fundamental forces of nature are (in order of decreasing strength): the strong force, electromagnetism, the weak force, and the force of gravity. Table 1.1 lists the forces with their relative strengths and the properties of the gauge bosons which mediate the interactions.

The strong force acts only on particles carrying the colour quantum number and is mediated by eight massless gluons. The strong force is described by the theory of Quantum Chromodynamics or QCD [2]. The familiar electromagnetic force acts on charged particles and is mediated by the photon while the weak force acts on quarks and leptons and is mediated by the $W^{\pm}$ bosons in the case of charged currents, and the Z^0 boson in the case of neutral currents. The electromagnetic and weak forces have been combined into a unified electro-weak theory [1]. Gravity has not been integrated into the Standard Model, although it is understood separately in the context of the theory of general relativity. Gravity is generally too weak on the distance scales of interest in high energy particle physics to be of importance in the calculations.

1.1.2 The Particles

The quarks and leptons of the Standard Model are divided into three families that exhibit identical properties with respect to the symmetries of the Standard Model except for their masses. Quarks carry a colour quantum number and hence experience the strong force whereas leptons do not. Six flavours of quarks and leptons exist. The quarks, in increasing order of mass, are the up (u), down (d), strange (s), charmed (c), bottom (b) and top (t). All of these quarks have been experimentally observed. The leptons are the electron (e), electron neutrino (ν_e), muon (μ), muon neutrino (ν_μ), tau (τ), and tau neutrino (ν_τ). All of the leptons, except for the tau neutrino, have been observed. The measured masses of the neutrinos are consistent with zero, although it is not expected that they are strictly equal to zero. Indirect cosmological arguments indicate that the sum of all of the neutrino masses is less than about $100\text{eV}/c^2$ [3].

The quarks and leptons are shown in equation 1.1 arranged into the three families of the Standard Model with the quarks in the top row and the leptons in the bottom row. The three "up type" quarks (u, c, and t) have electric charge $+2/3e$, while the three "down type" quarks (d, s, and b) have charge $-1/3e$. The relevant properties

Particle	Mass ($/c^2$)	Charge (e)
d	0.010 ± 0.005 GeV	-1/3
u	0.005 ± 0.003 GeV	+2/3
s	0.2 ± 0.1 GeV	-1/3
c	1.5 ± 0.2 GeV	+2/3
b	5.0 ± 0.3 GeV	-1/3
t	171 ± 10 GeV	+2/3
e	0.511 MeV	-1
μ	105.6 MeV	-1
τ	1784.1 MeV	-1
ν_e	< 7.3 eV	0
ν_μ	< 0.27 MeV	0
ν_τ	< 35 MeV	0

Table 1.2: Properties of the Standard Model Fermions

of all of the fermions are shown in Table 1.2 [4].

$$\begin{pmatrix} u \\ d \end{pmatrix} \quad \begin{pmatrix} c \\ s \end{pmatrix} \quad \begin{pmatrix} t \\ b \end{pmatrix}$$
$$\begin{pmatrix} e \\ \nu_e \end{pmatrix} \quad \begin{pmatrix} \mu \\ \nu_\mu \end{pmatrix} \quad \begin{pmatrix} \tau \\ \nu_\tau \end{pmatrix} \tag{1.1}$$

1.1.3 The Mathematical Framework

The Standard Model is expressed in the language of relativistic quantum gauge field theory. In practice this means that, for each of the forces except gravity, Lagrangian Quantum Field Theories exist in which all observables are finite. Of particular importance are the group symmetries observed by the theory. The electroweak interactions are described by a $SU(2)_L \times U(1)_Y$ symmetry while the strong interaction is described by the $SU(3)_{colour}$ structure of QCD. The combination of these symmetries into $SU(3)_{colour} \times SU(2)_L \times U(1)_Y$ gives the form of the Standard Model as a whole.

Within the Standard Model, particles exist as singlets, or as fundamental representations of the symmetry group inherent in the theory. Thus, in SU(2) particles

may be singlets or doublets, while in the SU(3) symmetry of QCD, particles exist as singlets or triplets. The gauge bosons, on the other hand, are adjoint to the fundamental representations and so are equal in number to the number of generators for the group. Thus we see how the gauge bosons of section 1.1.1 are generated from the theory. The single generator of U(1) translates into the single gauge boson (the photon) of electromagnetism, the three generators of SU(2) give the three bosons (W^+, W^-, and Z^0) of the weak interaction, and the eight generators of SU(3) correspond to the eight gluons.

The exact $SU(3)_{\text{colour}} \times SU(2)_L \times U(1)_Y$ symmetry of the Standard Model means that mass terms for the fermions and bosons are not present in the basic Lagrangian. Attempting to add mass terms by hand for the fermions or bosons causes the gauge invariance of the Lagrangian to be lost. This can be seen explicitly in the boson Lagrangian of the electroweak theory where the $SU(2)_L \times U(1)_Y$ symmetry gives rise to three gauge bosons from SU(2), $W_\mu^{1,2,3}$, and one from U(1), B_μ. If we define:

$$W^i_{\mu\nu} = \partial_\nu W^i_\mu - \partial_\mu W^i_\nu + g\epsilon^{ijk} W^j_\mu W^k_\nu, i = 1,2,3$$

and

$$B_{\mu\nu} = \partial_\nu B_\mu - \partial_\mu B_\nu$$

then the kinetic energy terms of the boson Lagrangian are given by:

$$\mathcal{L}_{KE} = -\frac{1}{4} W^i_{\mu\nu} W^{\mu\nu i} - \frac{1}{4} B_{\mu\nu} B^{\mu\nu}. \tag{1.2}$$

If explicit mass terms for the gauge bosons, such as $\frac{1}{2} m_B^2 B^\mu B_\mu$ are now introduced into this Lagrangian by hand, the invariance of the Lagrangian under the gauge transformation $B^\mu \rightarrow B^{\mu\prime} = B^\mu - \partial^\mu \chi / g$ and the theory loses its renormalization which gives rise to infinities in physical observables.

The problem of introducing masses into the theory can be dealt with in a subtle way through the *Higgs Mechanism* [5]. This mechanism, which is described in the following sections, makes use of Spontaneous Symmetry Breaking (SSB) to break the "exact" $SU(3)_{\text{colour}} \times SU(2)_L \times U(1)_Y$ symmetry of the Standard Model in such a

way as to introduce masses naturally into the theory. The process of applying SSB to the Standard Model gives rise to an observable scalar particle, the Higgs Boson. The experimental detection of this particle is the goal of this thesis. It should be stressed that while the Higgs Mechanism is a technically sufficient solution to the mass problem in the Standard Model, the physical interpretation of the Higgs sector is somewhat less well understood. Many believe that the detection of a Higgs Boson is the central experimental question facing particle physics today [6].

1.2 The Higgs Mechanism in the Standard Model

Within the Standard Model, the Higgs Mechanism is implemented through the introduction of a SU(2) doublet of two complex scalar Higgs fields

$$\phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} \tag{1.3}$$

with $\phi^+ = \frac{\phi_1 + i\phi_2}{\sqrt{2}}$ and $\phi^0 = \frac{\phi_3 + i\phi_4}{\sqrt{2}}$. A general SU(2) invariant potential is then associated with the Higgs doublet:

$$V = \mu^2 |\phi^\dagger \phi| + \lambda |\phi^\dagger \phi|^2 \tag{1.4}$$

where $\lambda > 0$ in order to bound the potential from below. If we choose $\mu^2 > 0$, then the minimum of the potential is at $\phi = 0$ and the symmetry is unbroken. For $\mu^2 < 0$, however, the potential takes the shape shown in Figure 1.1 and the minimum is no longer at $\phi = 0$ but instead is represented by a ring in the ϕ plane given by:

$$\phi^\dagger \phi = \frac{-\mu^2}{2\lambda} \equiv \frac{v^2}{2}.$$

Thus, no direction is preferred for the minimum, and we can choose a specific vacuum, ϕ_0, for the Higgs doublet.

$$\phi_0 = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v \end{pmatrix} \tag{1.5}$$

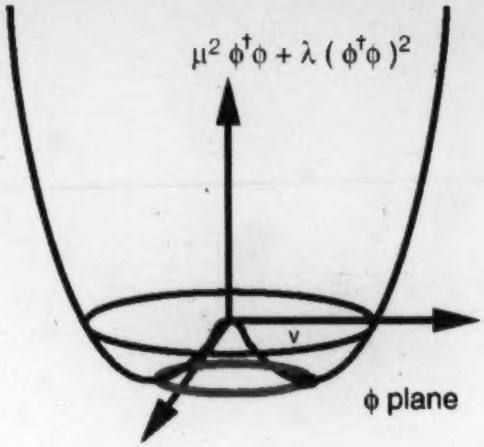


Figure 1.1: The Higgs potential. The minimum of the potential (green circle) has no preferred direction.

This choice gives a $U(1)_Y$ hypercharge of Y=1 for the Higgs doublet. We can now see the generation of masses for the gauge bosons by calculating the contribution to the Lagrangian from our Higgs scalar doublet. This contribution is given by:

$$\mathcal{L}_s = (D^\mu\phi)^\dagger(D_\mu\phi) - \mu^2|\phi^\dagger\phi| - \lambda|\phi^\dagger\phi|^2 \tag{1.6}$$

where the covariant derivative is given by $D_\mu = \partial_\mu - ig_1\frac{Y}{2}B_\mu - ig_2\frac{\vec{\tau}}{2}\cdot\vec{W}^\mu$ and g_1 and g_2 are respectively the $U(1)_Y$ and $SU(2)_L$ coupling constants. We can now expand our Higgs doublet around the minimum:

$$\phi = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 \\ v + h^0 \end{pmatrix}.$$

Substituting this expansion into our Lagrangian (equation 1.6) and setting Y=1 as we saw was the case from our choice of vacuum gives rise to the following terms in the kinetic energy portion of the Lagrangian:

$$\mathcal{L}_{KE} = \frac{1}{8}\left|\begin{pmatrix} g_1B_\mu + g_2W_\mu^3 & g_2(W_\mu^3 - iW_\mu^3) \\ g_2(W_\mu^1 + iW_\mu^2) & g_1B_\mu - g_2W_\mu^3 \end{pmatrix}\right|^2$$

$$= \frac{1}{8}v^2 g_2^2((W_\mu^1)^2 + (W_\mu^2)^2) + \frac{1}{8}v^2(g_1 B_\mu - g_2 W_\mu^3)^2. \quad (1.7)$$

As we will see in the next section, equation 1.7 will allow us to write mass terms for the gauge bosons.

1.2.1 The Gauge Boson Masses

We can rewrite the electroweak gauge bosons as physical gauge fields by re-writing equation 1.2 in terms of mass eigenstates. The four electroweak gauge bosons are then replaced by four physical gauge bosons: $W^\pm$, Z^0, and γ. These are defined as combinations of W_μ^{123} and B_μ as follows:

$$W^\pm{}_\mu = \frac{1}{\sqrt{2}}(W_\mu^1 \mp iW_\mu^2), Z^\mu = \frac{-g_1 B_\mu + g_2 W_\mu^3}{\sqrt{g_1^2 + g_2^2}}, A^\mu = \frac{g_2 B_\mu + g_1 W_\mu^3}{\sqrt{g_1^2 + g_2^2}}.$$

Re-writing the first term of equation 1.7 in terms of these mass states gives a mass term for the charged $W^\pm$ bosons equal to

$$\left(\frac{1}{2}vg_2\right)^2 W^+{}_\mu W^{-\mu} \quad (1.8)$$

while the second term of equation 1.7 gives the Z^0 boson mass term:

$$\frac{1}{4}v^2\left(g_1^2 + g_2^2\right)^2 Z_\mu Z^\mu. \quad (1.9)$$

As there is no mass term generated for the A_μ field, the following masses are generated for the physical gauge bosons through the Higgs Mechanism:

$$m_{W^\pm} = \frac{1}{2}vg_2$$

$$m_{Z^0} = \frac{1}{2}v(g_1^2 + g_2^2)$$

and

$$m_\gamma = 0.$$

Prior to the spontaneous symmetry breaking, a total of twelve degrees of freedom were present in the boson system: four from the complex scalar Higgs doublet, six

from the massless W^i_μ, and two from the massless B_μ field. After SSB, we have the same number of degrees of freedom: Nine from the massive $W^\pm$ and Z^0 bosons, two from the massless photon, and one from our new scalar Higgs particle, h^0. It is common in the context of the Higgs sector to state that three of the scalar degrees of freedom present before SSB have been *eaten* to form the longitudinal components of the W^+, W^-, and Z^0.

1.2.2 The Fermion Masses

The Higgs Mechanism also allows fermion masses to be introduced naturally into the theory. We can add a Yukawa lepton interaction term to our previous Lagrangian,

$$\mathcal{L}_{\text{int}} = g_e \left(\overline{L} \phi \bar{e_R} + \phi^\dagger \overline{\bar{e_R}} L \right) \tag{1.10}$$

where g_e is an arbitrary coupling, $\phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix}$ is our Higgs doublet, and the left-handed lepton doublet is defined by $L_{L_k} = \begin{pmatrix} \nu_k \\ e_k \end{pmatrix}$. It may be seen that $\overline{L}\phi = \overline{\nu}_{eL}\phi^+ + \overline{\bar{e_L}}\phi^0$ is SU(2) invariant and subsequently multiplying by the $\bar{e_R}$ singlet does not destroy this invariance. The second term in equation 1.10 is simply the Hermitian conjugate of the first term.

We can see that the overt addition of mass terms of the form $m_e \bar{e} e = m_e(\overline{\bar{e_L}} \bar{e_R} + \overline{\bar{e_R}} \bar{e_L})$, to the Lagrangian is forbidden by gauge invariance as the left-handed and right-handed fermions transform differently under SU(2) and U(1). However, if we allow the Higgs doublet to take on its vacuum expectation value:

$$\phi \rightarrow \begin{pmatrix} 0 \\ \frac{v+h}{\sqrt{2}} \end{pmatrix} \tag{1.11}$$

we have, substituting into equation 1.10:

$$\mathcal{L}_{\text{int}} = \frac{g_e}{\sqrt{2}} \left(\overline{\bar{e_L}} \bar{e_R} + \overline{\bar{e_R}} \bar{e_L} \right) + \frac{g_e}{\sqrt{2}} \left(\overline{\bar{e_L}} \bar{e_R} + \overline{\bar{e_R}} \bar{e_L} \right) h \tag{1.12}$$

where h is the neutral physical Higgs particle.

Now, the first term of equation 1.12 has exactly the form of a mass term for a fermion. Hence, we can write:

$$m_e = \frac{g_e v}{\sqrt{2}} \tag{1.13}$$

and thus the theory can *accomodate* a non-zero mass for the electron. We say *accomodate* since the arbitrary coupling g_e does not allow us to predict the mass of the electron. No mass term for the neutrino ocurred in equation 1.12 as a result of the fact that the theory is assumed to contain no right-handed neutrino ν_R which could be used in the interaction Lagrangian (equation 1.10). This implies that neutrinos do not interact with the Higgs boson. The same calculation can be performed for the other lepton families with similar results.

The situation for quarks is a little more complicated due to the fact that right-handed states exist for all flavours of quarks. It is well known in ordinary spin theory that if $\psi = \begin{pmatrix} a \\ b \end{pmatrix}$ is an SU(2) doublet, then so is:

$$\psi_c = -i\tau_2\psi^* = \begin{pmatrix} -b^* \\ a^* \end{pmatrix}.$$

We can then write terms in our interaction Lagrangian using this fact of the form:

$$\phi_c = \begin{pmatrix} -\phi^0 \\ \phi^- \end{pmatrix}. \tag{1.14}$$

Since we know that ϕ has U(1) hypercharge $Y = +1$, ϕ_c has hypercharge $Y = -1$, and still satisfies $Q = T_3 + Y/2$. We can then write an interaction Lagrangian for the quarks:

$$\mathcal{L}_{\text{int}} = g_d \overline{Q}_L \phi d_R + g_u \overline{Q}_L \phi_c u_R + h.c. \tag{1.15}$$

where g_u and g_d are arbitrary Yukawa couplings. By substituting from equations 1.11 and 1.14, and using $Q_{L\alpha} = \begin{pmatrix} u_\alpha \\ d_\alpha \end{pmatrix}$ (α is a color index), we obtain terms of the type:

$$\frac{v+h}{\sqrt{2}} g_d \bar{d}d \tag{1.16}$$

from which the mass term

$$m_d = \frac{g_d v}{\sqrt{2}} \tag{1.17}$$

arises. This procedure may be applied to all fermions to give the general mass term

$$m_f = \frac{g_f v}{\sqrt{2}}. \tag{1.18}$$

As for leptons, this procedure may be extended to all families of quarks by changing the Yukawa couplings into $N_F \times N_F$ matrices where F is the number of families. As we saw in equation 1.18, the mass matrices and the Yukawa matrices are proportional for fermions, and hence the Higgs does not mediate any flavour changing interactions.

1.2.3 Properties of the Higgs Boson

In order to search for a Higgs Boson experimentally, we must investigate some of its properties. From equation 1.16, we see that the coupling between the Higgs and fermions is given by:

$$g_{Hf} = \frac{g_f}{\sqrt{2}} = \frac{m_f}{v} = \frac{g_2}{2M_W} m_f. \tag{1.19}$$

This means that the Higgs coupling to fermions is proportional to the fermion mass and the Higgs will preferentially decay to the heaviest fermions kinematically available. Similarly, the coupling of the Higgs to other gauge bosons may be calculated to be of the form $g_W = g_2 M_W$.

The branching ratio of the Higgs boson into fermions is proportional to the fermion mass squared and is given by:

$$BR(h^0 \to f\bar{f}) = \frac{N_c m_f^2 \beta_f^3}{\sum_{f'} N_c m_{f'}^2 \beta_{f'}^3} \tag{1.20}$$

where N_c is the number of colours, $\beta_f = \sqrt{1 - \frac{4m_f^2}{m_H^2}}$ and the sum is over all available channels.

Figures 1.2 and 1.3 [7] show the branching ratios for Higgs decay into fermions and bosons. For Higgs masses in the range of 10 GeV to 160 GeV (the W pair production threshold), $h \to b\bar{b}$ is the most important decay mode. After the W pair threshold, the branching ratios to fermions decrease as the $W^{\pm}$ branching ratio

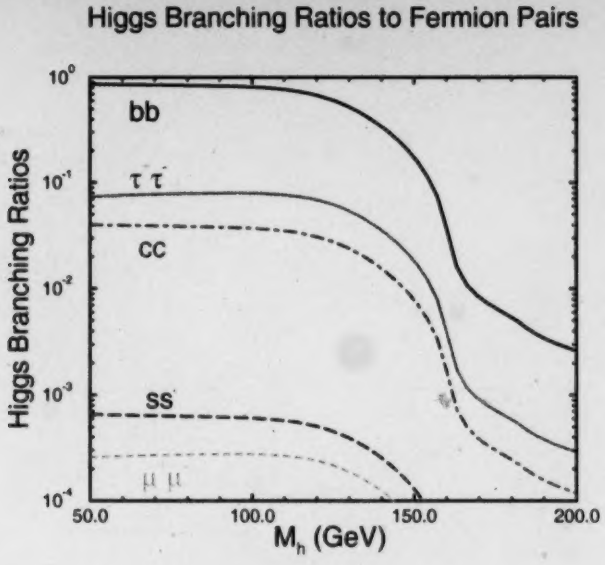


Figure 1.2: Branching ratios for Standard Model Higgs boson decay to fermions.

increases. For the kinematic region accessible at LEP2, however, Higgs bosons decay almost exclusively to $b\bar{b}$ pairs, although enough $\tau^+\tau^-$ decays occur to warrant investigation.

1.2.4 Production of Higgs Bosons at LEP

Since the coupling of the Higgs boson to the electron directly is small ($\sim m_e/v$), production of Higgs bosons through the s channel is suppressed, and the main production mechanism is the so called "Higgstrahlung" channel where the Higgs is produced together with a virtual (at LEP1) or real (at LEP2) Z^0 boson. This process, shown in figure 1.4 has a cross section [8] equal to:

$$\sigma(e^+e^- \to Z^0h^0) = \frac{\pi\alpha^2\lambda_{Z^0h^0}^{1/2}[\lambda_{Z^0h^0} + 12\frac{m_Z^2}{s}][1 + (1 - 4\sin^2\theta_W)^2]}{192s\sin^4\theta_W\cos^4\theta_W(1 - m_Z^2/s)^2} \tag{1.21}$$

where

$$\lambda_{Z^0h^0} \equiv (1 - \frac{m_h^2 + m_Z^2}{s})^2 - 4\frac{m_h^2 + m_Z^2}{s^2}.$$

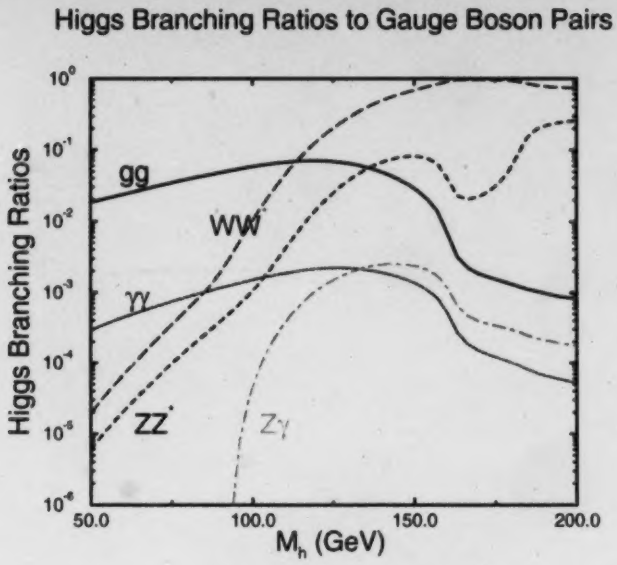


Figure 1.3: Branching ratios for Standard Model Higgs boson decay to bosons.

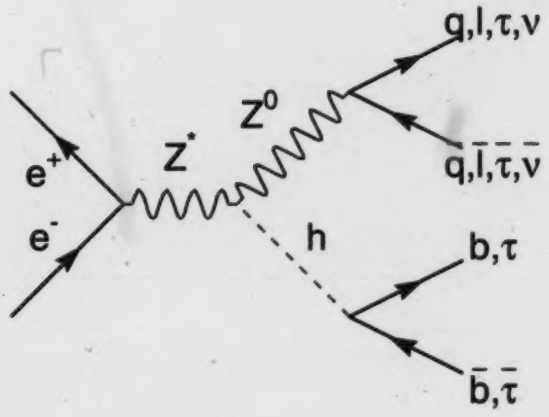


Figure 1.4: The Higgstrahlung production mechanism. Common decays of the Higgs boson and the Z^0 are also shown.

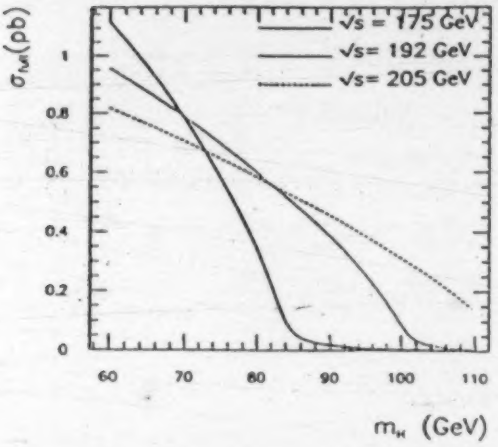


Figure 1.5: Cross section for Higgstrahlung production versus Higgs mass for various centre of mass energies.

This cross section is at a maximum for $\sqrt{s} = m_Z + 2m_h$ and increases quickly with increasing beam energy. Hence, one wants to search for the Higgs at the highest beam energy available. Figure 1.5 [7] shows the cross sections for Higgstrahlung production for various centre of mass energies versus the Higgs mass.

Higgs production also proceeds through vector boson fusion at LEP2, but this mechanism is suppressed by an additional electroweak vertex, and hence proceeds at a much lower rate than the Higgstrahlung procedure. This mechanism is not considered further in this thesis as it does not contribute significantly to the production rate.

The high proportion of b quarks in Higgs decays means that any successful analysis will have to make use of effective b quark tagging. The main backgrounds to the Higgs signal at LEP2 will be described more fully in chapter 4, and consist of four fermion final state events from $Z^0 \to q\bar{q}(\gamma)$, $W^+W^- \to q\bar{q}q\bar{q}$, and $Z^0Z^0 \to q\bar{q}q\bar{q}$.

Analysis of OPAL data from past LEP2 running at $\sqrt{s} = 130 - 172$ GeV has allowed a lower limit on the mass of a Standard Model Higgs boson to be set at

69.4 GeV at the 95% confidence level [9].

1.3 Successes and Deficiencies of the Standard Model

The main tenets of the Standard Model have now been presented. The theory is represented by a $SU(3)_c \times SU(2)_L \times U(1)_Y$ gauge group with one Higgs doublet. The fermions are represented by left-handed weak isodoublets and right-handed singlets. There are 21 free parameters in the theory. These are:

- 3 coupling constants.
- 12 fermion masses.
- 4 fermion mixing paramters.
- 1 Higgs boson mass.
- 1 independent gauge boson mass.

The Standard Model has had many successes to date including the unification of the electromagnetic and weak forces into a single renormalizable theory, prediction of the W and Z masses, prediction of the structure of the weak charged and neutral currents, and agreement with all available experimental data [10].

Despite these successes, there are a number of important questions which are not addressed by the Standard Model in its current state, and it is for these reasons that many believe the Standard Model to be only a subset of a more global theory. These questions include:

- Why are the fermion masses and mixings not predicted?
- Are there massive neutrinos?
- Why do there seem to be three particle generations?

- Why is the charge of the electron exactly opposite the charge of the proton?
- Why is there no dark matter candidate in the theory?
- How does gravity fit into the model?
- Why are there so many free parameters in the theory?

Many attempts at answering these questions have been made in the course of extending the Standard Model. With respect to the Higgs boson, one of the most perplexing problems is that radiative corrections to the Higgs mass may be many times the mass of the particle itself. This *gauge hierarchy problem* and a possible solution in the form of the Minimal Supersymmetric Model, is presented in the next section.

1.4 The Minimal Supersymmetric Model

The Minimal Supersymmetric Model (MSSM) [11] is an extension of the Standard Model that attempts to relate fermions and bosons. Particles of different spin are united into "superfields" with each scalar having a fermionic partner and each fermion having a scalar partner. Pairs of partner particles share all quantum numbers exactly except for spin.

The MSSM is called an *extension* to the Standard Model as it respects the same $SU(3)_c \times SU(2)_L \times U(1)_Y$ symmetry as the Standard Model and attempts to introduce supersymmetry with a minimum of disturbance to the predictive power of the SM. Since no experimental candidates for the supersymmetric partners of known particles have been found, the MSSM introduces a spectrum of "sparticles" [12] in effect doubling the number of particles in the theory.

From the point of view of the Higgs sector and the gauge hierarchy problem [13] mentioned in section 1.3, the MSSM is particularly appealing as the way the scalars and fermions in the MSSM couple to gauge bosons allows the problem to be rectified.

This solution to the gauge hierarchy problem will serve as a motivation for studying the Higgs sector of the MSSM.

1.4.1 Motivation: The Gauge Hierarchy Problem

Fits to electroweak data from LEP, SLD, and the Tevatron show a minimum χ^2 near a value of 100 GeV/c^2 for the Higgs mass (figure 1.6) [10]. This estimate, however, does not take into account corrections due to radiation of virtual particles. For a fermion with mass m_f the coupling with Higgs boson gives rise to a term

$$-g_f h^0 f\bar{f}$$

in the Lagrangian which results in a correction to the Higgs mass [14] given by:

$$\Delta m_h^2 = N_c \frac{g_f^2}{16\pi^2}\left[-2\Lambda^2 + 6m_f^2 \ln(\Lambda/m_f) + \ldots\right] \tag{1.22}$$

where Λ is the *ultraviolet momentum cutoff* that defines the energy scale at which new physics will enter the model and N_c is the number of colours (3 for quarks, 1 for leptons). It can be seen from equation 1.22 that the mass corrections to the Higgs diverge quadratically with Λ and that this behaviour does not depend on the fermion mass. Thus, for an ultraviolet cutoff equal to the Plank scale at $m_p = 1.2 \times 10^{19}$ GeV/c^2, the corrections to m_h turn out to be thirty orders of magnitude greater than the $m_h \sim \mathcal{O}(100$ GeV/c$^2)$ original Higgs mass. This is also a problem for the masses of other particles since, as we saw above, the masses of all fermions and bosons depend on the Higgs.

This quadratic divergence in Λ is also found in the interaction of the SM Higgs boson with heavy complex scalars. If such a scalar S were to exist, it would interact with the Higgs with a term $-g_S h^0 h^0 SS$ in the Lagrangian. This would also give a radiative mass correction to the Higgs equal to

$$\Delta m_h^2 = \frac{g_S}{16\pi^2}\left[\Lambda^2 - 2m_s^2 \ln(\Lambda/m_S) + \ldots\right]. \tag{1.23}$$

The relative minus sign between the Λ^2 terms in equations 1.22 and 1.23 suggests a way in which to neatly cancel the mass correction divergence without having to

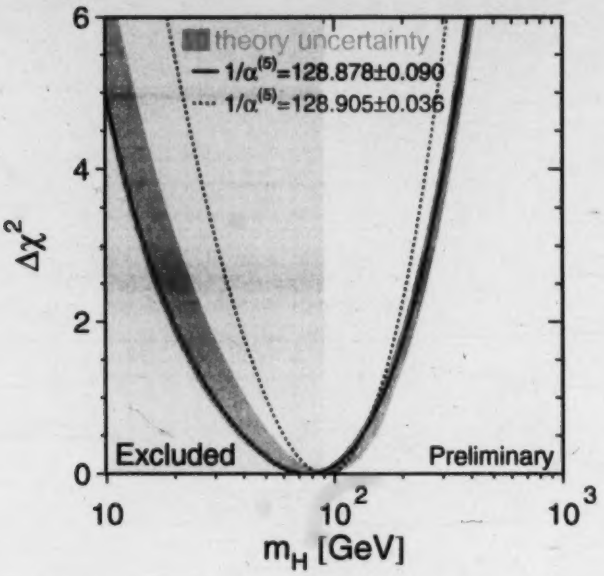


Figure 1.6: Estimate of the Higgs mass from fits to electroweak data. The χ^2 of the fit favours a Higgs mass of about 100 GeV/c^2.

reduce the scale at which new physics is expected in the model. If each of the Standard Model fermions f was accompanied by two complex scalars S with $g_S = |g_f|^2$, then the diverging terms in Λ would cancel between the two equations. It is through the application of a supersymmetry between the fermions and bosons that this cancellation is achieved.

1.4.2 The Higgs Mechanism in the MSSM

In the MSSM, two scalar Higgs doublets are used to generate the spontaneous symmetry breaking necessary to bring masses into the theory. The first doublet, ϕ_1 couples to up type quarks and the second, ϕ_2, couples to down type quarks and charged leptons. We can write a general MSSM scalar potential [6] as

$$\begin{aligned} V_H &= \left(|\mu|^2 + m_1^2\right)|\phi_1|^2 + \left(|\mu|^2 + m_2^2\right)|\phi_2|^2 - \mu B \epsilon_{ij}\left(\phi_1^i \phi_2^j + \text{h.c.}\right) \\ &\quad + \frac{g_1^2 + g_2^2}{8}\left(|\phi_1|^2 - |\phi_2|^2\right)^2 + \frac{1}{2}g_2^2 |\phi_1^* \phi_2|^2 \end{aligned} \tag{1.24}$$

where we have three independent combinations of parameters: $|\mu|^2 + m_1^2$, $|\mu|^2 + m_2^2$, and μB.

Spontaneous symmetry breaking occurs in the MSSM in much the same way that it does in the Standard Model. It may be readily seen that all three terms in the potential of equation 1.24 must be non-zero in order to move the minimum of the potential away from $V_H = 0$. We then choose vacuum expectation values for both of the Higgs doublets:

$$\phi_1^0 = \begin{pmatrix} v_1 \\ 0 \end{pmatrix}, \phi_2^0 = \begin{pmatrix} 0 \\ v_2 \end{pmatrix} \tag{1.25}$$

where v_1 and v_2 are chosen to be positive.

By substituting equation 1.25 into the potential, the W gauge boson gets a mass determined by v_1 and v_2 in a manner analagous to the Standard Model mechanism

$$m_{W^\pm}^2 = \frac{g_2^2}{2}(v_1^2 + v_2^2).$$

The known experimental mass of the $W^\pm$ is then used to fix the value of $v_1^2 + v_2^2$.

In a manner similar to the SM analysis, we can now compare degrees of freedom in the system prior to and after spontaneous symmetry breaking. Prior to SSB, there were eight degrees of freedom from the 2 Higgs doublets. As in the SM, three of these are "eaten" to give mass to the $W^\pm$ and Z^0 bosons, and we are left with five degrees of freedom instead of the one present in the Standard Model. This implies the existence of a five particle Higgs sector in the MSSM consisting of a charged Higgs boson, $H^\pm$, a CP-odd neutral Higgs boson, A^0, and two CP-even neutral Higgs bosons, h^0 and H^0. By convention, h^0 is taken to be the less massive of the two CP-even neutral Higgs bosons.

It can be shown that the masses of the physical Higgs bosons [6] may all be found in terms of two parameters. The standard choices for these parameters are:

$$\tan\beta \equiv \frac{v_2}{v_1} \tag{1.26}$$

and the mass of the CP-odd neutral Higgs boson, m_{A^0}. These masses are found by diagonalizing the 2×2 Higgs mass matrix and thereby introduces the mixing angle,

α between the neutral Higgs bosons in the MSSM as

$$\tan 2\alpha = \frac{\sin 2\beta(m_A^2 + m_Z^2)}{\cos 2\beta(m_A^2 - m_Z^2)}. \tag{1.27}$$

At tree level, the masses of the neutral MSSM Higgs bosons are given by

$$m_{h,H}^2 = \frac{1}{2}\left[m_h^2 + m_Z^2 \mp \left((m_A^2 + m_Z^2)^2 - 4m_Z^2 m_A^2 \cos^2 2\beta\right)^{1/2}\right],$$

the CP-odd boson mass is determined by

$$m_A^2 = \frac{2|\mu B|}{\sin 2\beta},$$

and the charged Higgs masses are given by

$$m_{H^\pm}^2 = m_{W^\pm}^2 + m_A^2.$$

These equations would seem to imply that the lightest MSSM Higgs boson, h^0, is lighter than the Z^0 boson, but corrections to the mass from loops of top quarks and top squarks are large. At one loop, the neutral Higgs boson masses become

$$\begin{aligned} m_{h,H}^2 \;=\; & \frac{1}{2}\Big\{m_{A^0}{}^2 + m_{Z^0}{}^2 + \frac{\epsilon_h}{\sin^2\beta} \pm \Big[\Big((m_{A^0}{}^2 - m_{Z^0}{}^2)\cos 2\beta + \frac{\epsilon_h}{\sin^2\beta}\Big)^2 \\ & + \Big(m_{A^0}{}^2 + m_{Z^0}{}^2\Big)^2 \sin^2 2\beta\Big]^{1/2}\Big\} \end{aligned} \tag{1.28}$$

where the contribution from one loop [6] is given by

$$\epsilon_h \equiv \frac{3G_F}{\sqrt{2}\pi^2} m_t^4 \ln\left(1 + \frac{m_{\tilde{t}}^2}{m_t^2}\right).$$

We can see that the corrections at one loop are always positive and the mass of h^0 varies with the increasing top quark mass. The resulting shift in m_{h^0} is approximately proportional to

$$m_t^4 \ln\left(\frac{m_{\tilde{t}}}{m_t^2}\right)$$

which is on the order of tens of GeV for the measured top mass of 175 GeV/c^2.

The couplings of the MSSM Higgs bosons are conventionally stated in terms of the Standard Model couplings

$$\mathcal{L} = \frac{g_2}{2m_W}\left[c_{ffh}\overline{f}_i f_i h^0 + C_{ffH}\overline{f}_i f_i H^0 + C_{ffA}\overline{f}_i \gamma_5 f_i A^0\right] \tag{1.29}$$

f	C_{ffh}	C_{ffH}	C_{ffA}
u	$\frac{\cos\alpha}{\sin\beta}$	$\frac{\sin\alpha}{\sin\beta}$	$\cot\beta$
d	$-\frac{\sin\alpha}{\cos\beta}$	$\frac{\cos\alpha}{\cos\beta}$	$\tan\beta$

Table 1.3: Coupling constants of the MSSM Higgs bosons to fermions.

where $C_{ffh} = 1$ for a Standard Model Higgs boson. For fermions, the couplings are no longer proportional to mass, but depend on both the fermion masses and the vacuum expectation values of the neutral components of the Higgs doublets. The couplings depend on both β (equation 1.26) and α (equation 1.27). Table 1.3 gives the couplings in terms of α and β for up and down type quarks. For values of $\tan\beta > 0.7$, the decay properties h^0 are nearly identical to the Standard Model Higgs.

Couplings of the neutral scalar MSSM Higgs bosons to vector bosons are suppressed relative to the equivalent Standard Model couplings while no tree level coupling between A^0 and the vector bosons exists. Complete listings of the MSSM Higgs sector coupling constants may be found elsewhere [6].

The predominant mechanism for creation of the neutral scalar Higgs bosons at LEP is still through the Higgstrahlung process. A second process, the so called *pair production* process $e^+e^- \rightarrow h^0A^0$ also is of significance at LEP. In fact, the two production mechanisms are complementary as can be seen by relative factors of $\cos^2(\beta-\alpha)$ and $\sin^2(\beta-\alpha)$ in their cross sections:

$$e^+e^- \rightarrow h^0Z^0 : \sigma_{hZ} = \sin^2(\beta-\alpha)\sigma_{HZ}^{SM} \tag{1.30}$$

$$e^+e^- \rightarrow h^0A^0 : \sigma_{hA} = \cos^2(\beta-\alpha)\overline{\lambda}\sigma_{\overline{\nu}\nu}^{SM} \tag{1.31}$$

where σ_{HZ}^{SM} and $\sigma_{\overline{\nu}\nu}^{SM}$ are the Standard Model cross sections for the processes $e^+e^- \rightarrow h_{SM}^0Z^0$ and $e^+e^- \rightarrow \overline{\nu}\nu$ respectively and $\overline{\lambda}$ is a kinematic factor. The fact that equation 1.30 gives the MSSM Higgstrahlung cross section in terms of the Standard Model cross section, and the fact that the decay properties of h^0 closely match h_{SM}^0 for $\tan\beta > 0.7$ allows easy transfer of analyses between the search for the Standard

Model Higgs boson and the lightest MSSM neutral scalar Higgs boson, a fact that will be exploited in chapter 4.

1.4.3 The MSSM Scan at OPAL

The most general MSSM has over one hundred tuneable parameters. A constrained theory [15] is considered in the present analysis that contains only six additional free parameters as compared to the Standard Model. This constrained model assumes unification of

- The scalar-fermion masses, m_0 at the grand unification scale.
- The gaugino masses parameterized by M_2.
- The SU(2) gaugino mass term at the electroweak scale.
- The scalar-fermion tri-linear couplings (A) at the electroweak scale.

and therefore has m_0, M_2, A, the Higgs mass parameter μ, $\tan\beta$, and the mass of the CP-odd neutral Higgs m_{A^0} as free parameters. This constrained MSSM has virtually no effect on the Higgs sector as compared to a general MSSM.

In OPAL, a general MSSM scan is employed to map the performance of the many Higgs search channels in different regions of $m_{A^0} - \tan\beta$ space. Many of the search channels are complementary and allow large regions of the parameter space to be excluded without holes. Determining the cross-sections and branching ratios for every possible combination of MSSM parameters requires a huge amount of computing power, and for this reason the OPAL MSSM scan was performed centrally at CERN. Knowing the cross-sections and branching ratios at each point in parameter space allows the expected Higgs signal to be calculated for each point. This expected signal can then be compared to the experimental results, and a confidence level analysis [15] can be applied to exclude regions of the parameter space.

To perform the general scan, the values of the six free parameters presented above are varied independently in the Monte Carlo generators HZHA [16] and SUSYGEN

[17]. These generators provide the masses and couplings of all MSSM Higgs bosons as well as the supersymmetric particles in addition to the cross-sections for the Higgstrahlung and pair production mechanisms at a given centre of mass energy. The six free MSSM parameters in the constrained model, and the mass of the top quark, were selected randomly and independently from the following ranges[1]. The parameter ranges were divided into bins of varying size for the random choices in order to step logically through the parameter space. The total number of parameter combinations considered in the general MSSM scan was close to 6,000,000.

- m_0: 0 to 1000 GeV.
- M_2: 0 to 2000 GeV.
- A: $-2.5m_0$ to $2.5m_0$.
- μ: -1000 to 1000 GeV.
- m_{A^0}: 5 to 160 GeV.
- $\tan\beta$: 0.7 to 50.
- m_t: 165, 175, and 185 GeV. This represents two standard deviations from the measured top quark mass.

Exclusion plots for the general MSSM scan using data from LEP2 with $\sqrt{s} \leq$ 172 GeV are shown in figure 1.7 [15] .

[1] Except for the mass of the top quark for which three discrete values were used.

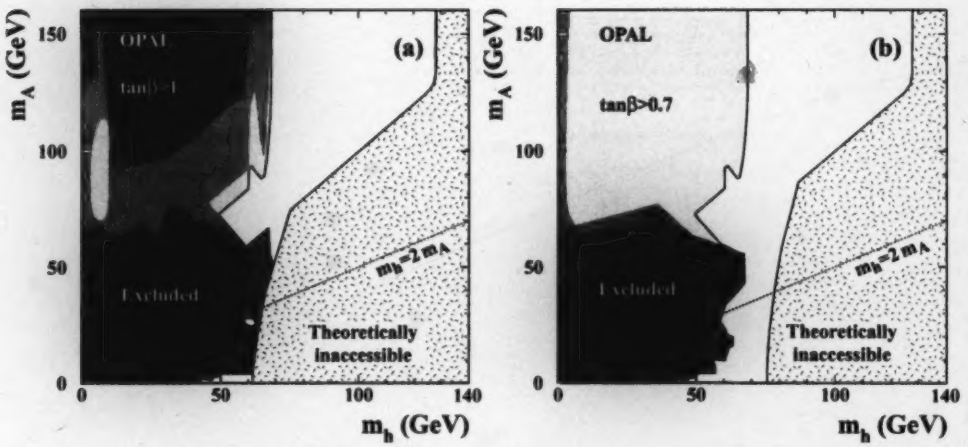


Figure 1.7: MSSM exclusion plots for $\sqrt{s} \leq 172$ GeV. Excluded regions are shown in the (m_{h^0} , m_{A^0}) plane at the 95% confidence level. The left plot shows the result for $\tan\beta > 1$ and the right plot for $\tan\beta > 0.7$. The black areas are excluded solely by comparing data to the expected MSSM production rates, without applying any additional theoretical criteria. The speckled grey areas are theoretically inaccessible.

Chapter 2

The Experimental Apparatus

The search for the Higgs Boson presented in this thesis was carried out using data collected by the Omni Purpose Apparatus for LEP (OPAL). This experiment is one of four large particle physics experiments collecting data from the Large Electron Positron (LEP) collider at the European Laboratory for Particle Physics (CERN[1]) near Geneva, Switzerland. CERN provides high energy particle physics research facilities to over half of the world's practicing particle physicists.

The main accelerator complex at CERN is shown in figure 2.1. The complex at CERN is centered on the Proton Synchrotron (PS) which was originally designed as a proton accelerator for fixed target experiments. The PS has been continuously upgraded since it was commissioned, and is now capable of accelerating electrons, protons, and heavy ions in order to feed the other accelerators at CERN. The Super Proton Synchrotron (SPS) is the accelerator that produced the first W and Z bosons as observed in the UA1 experiment [18]. It now accelerates protons and antiprotons for collider and fixed target experiments as well as accelerating electrons and positrons for injection into LEP. The Large Electron Positron collider itself is a ring 27 km in circumference that is now capable of creating electron positron collisions at a centre of mass energy of 200 GeV. The search for a Standard Model Higgs Boson presented in this thesis makes use of data recorded at a centre of mass energy of

[1]The acronym CERN refers to the French name for the Laboratory: Conseil Européen pour la Recherche Nucléaire.

183 GeV taken during 1997.

2.1 The Large Electron Positron Collider

The Large Electron Positron (LEP) collider was originally designed as a production factory for Z^0 bosons, and hence had a design energy of approximately 45 GeV per beam in order to produce collisions at the Z^0 resonance. LEP operated in this configuration for seven years after its commissioning in 1989. During this period, LEP was able to operate at or near its design luminosity for much of the time, and some 160 pb^{-1} of Z^0 decays were recorded by the OPAL experiment alone.

Starting in November of 1995, superconducting magnets were added to the LEP ring in order to allow centre of mass energies above the Z^0 resonance. These magnets were added in several stages from 1995 to 1998 resulting in runs being taken at several centre of mass energies over the past four years. Runs were taken at centre of mass energies of 130, 161, 172, and 183 GeV. Since the publication of the analysis presented in this thesis, runs have taken place at centre of mass energies of 189, 192, and 200 GeV. Electrons and positrons are brought to these high collision energies through a series of acceleration steps. The CERN accelerator complex that injects the particles into LEP is shown in figure 2.1. Electrons are first accelerated to 200 MeV in the electron linear accelerator (LIL) and directed onto a target in order to produce positrons. The positrons are then accelerated to an energy of 600 MeV in the second stage of LIL before being injected into the Electron-Positron Accumulator (EPA) where they are stored and compressed by synchrotron damping. This process is repeated until about 3×10^{11} positrons in each of eight bunches have been collected in the EPA, a goal which can take up to 1000 bursts from LIL to achieve. This beam is then injected into the Proton Synchrotron where it is ramped up to an energy of 3.5 GeV before being injected into the Super Proton Synchrotron (SPS) and further accelerated to 20 GeV. At this point, the positron beam is injected into LEP in a clockwise direction and ramped up to its final energy of 91.5 GeV for

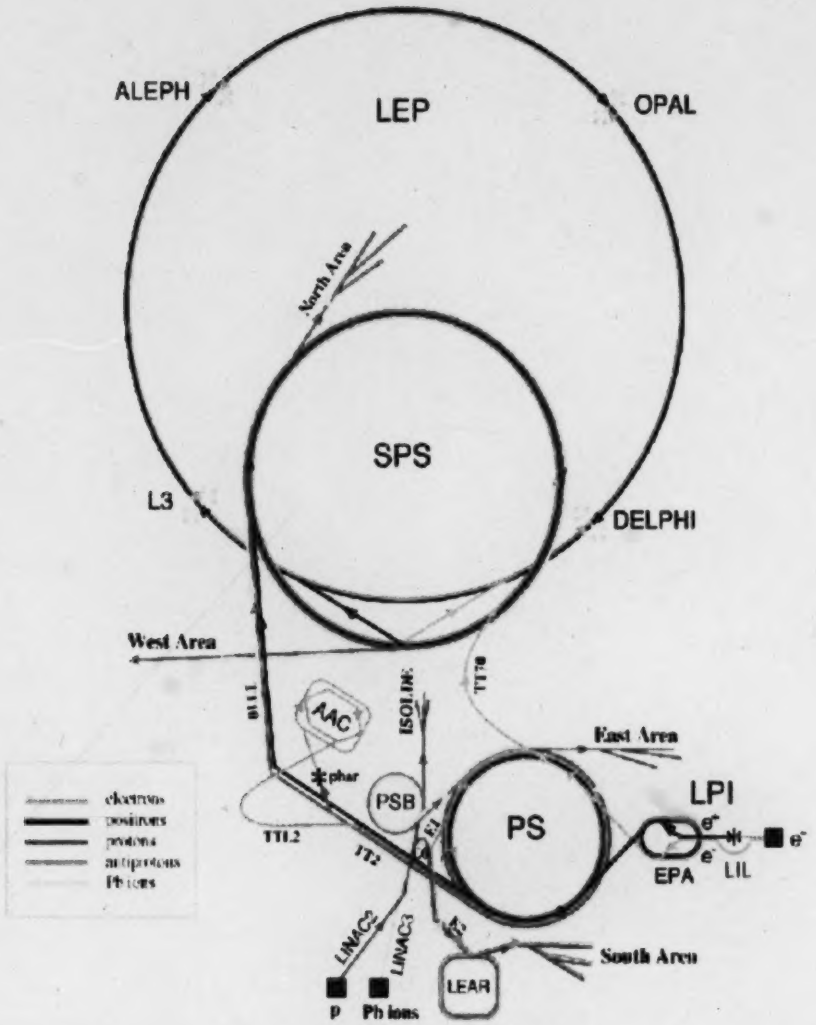


Figure 2.1: The CERN accelerator complex near Geneva, Switzerland.

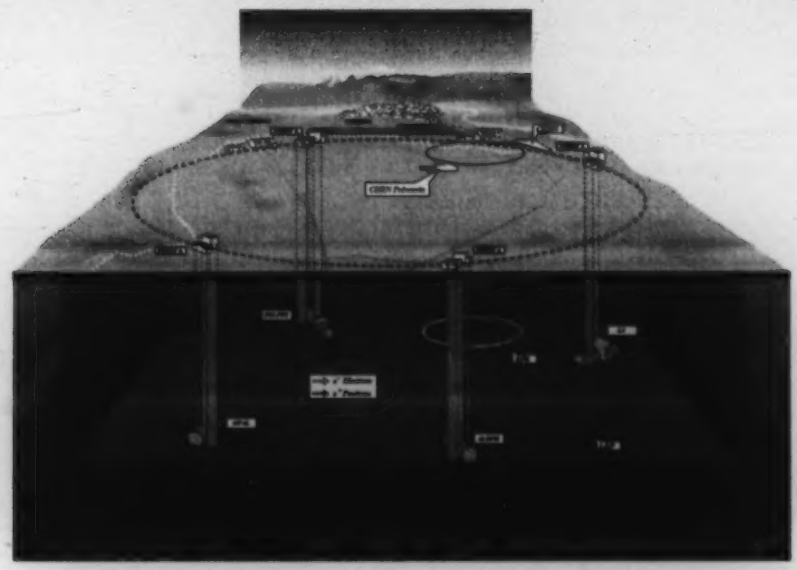


Figure 2.2: The four LEP experiments at the their respective interaction points. The OPAL detector is located at Interaction Point 6.

183 GeV running. The electrons are subjected to the same acceleration process as the positrons except that the LIL target is removed and the SPS is set to inject them into LEP in a counter-clockwise direction. LEP itself is divided into eight straight sections connected by 2.8 km long arcs. The straight sections contain both the detector halls and the RF accelerating cavities while the curved sections hold the dipole magnets for beam steering. Quadrupole magnets are situated at various positions throughout the LEP ring in order to adjust the focus of the beam. Collisions between the electron and positron beams take place at four instrumented Interaction Points (IP's) on the LEP ring. The four detectors are shown at their respective LEP interaction points in figure 2.2. Although collisions are possible at more than the four instrumented interaction points, electrostatic separators at these sites keep the beams separated in order to extend the beam lifetime. The average lifetime of the LEP beams, during which data taking is possible, is about six hours.

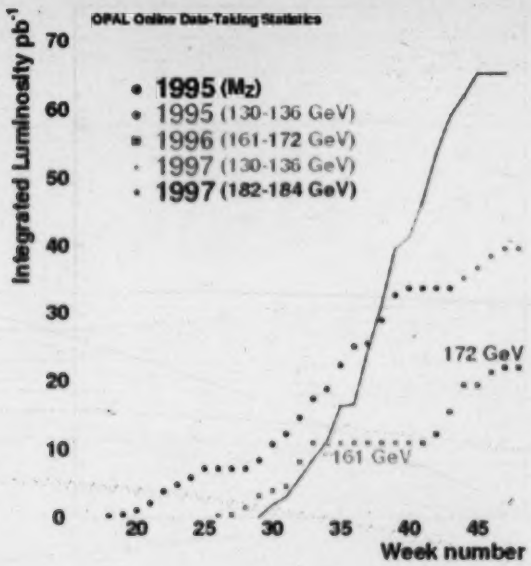


Figure 2.3: Integrated luminosity delivered to OPAL from 1991 to 1997.

The goal of LEP is to provide a high luminosity for the four detectors. Since the rate for a given reaction is obtained by multiplying the cross section of the reaction by the luminosity, the size of the data set accessible to physicists depends linearly upon the luminosity delivered by the LEP beams. Each of the LEP beams consists of eight bunches of particles, either electrons or positrons depending on the direction of circulation of the beam. At the interaction points, the bunches have a three dimensional gaussian shape with standard deviations of $\sigma_x \approx 200\,\mu$m, $\sigma_y \approx 4\,\mu$m, and $\sigma_z \approx 15$mm. The beam dimensions are important to physics running as the luminosity is inversely proportional to the effective cross section of the colliding beams. The luminosity $\mathcal{L}$ delivered by LEP is calculated through the relation:

$$\mathcal{L} = fn\frac{N_1 N_2}{4\pi\sigma_x\sigma_y}$$

where f is the frequency of the bunch revolution (11 kHz), n is the number of bunches, N_1 is the number of electrons per bunch, and N_2 is the number of positrons

per bunch. The 1997 run saw LEP deliver the best integrated luminosity [2] of its history to that point in time with over 65pb^{-1} being recorded by OPAL. This was due both to increased understanding of the beam optics and an improved duty cycle for the accelerator. The integrated luminosity delivered as a function of weeks is shown in figure 2.3 for the runs of 1991 through 1997. This thesis uses 54.1pb^{-1} of the data taken at $\sqrt{s} = 183$ GeV during the 1997 run.

2.2 The OPAL Detector

The OPAL (Omni Purpose Apparatus for LEP) detector is one of the four large particle physics detectors that collect data on the LEP ring. OPAL is a collaboration of over 150 physicists from ten countries (see Appendix A).

The OPAL detector [19], shown in figure 2.4, is cylindrical in shape and centered axially on the LEP beam pipe. The detector is inclined at an angle of 13.9 mrad with respect to the horizontal plane following the slight slope of LEP. The OPAL coordinate system, also shown in figure 2.4, is defined so that the z axis lies along the beam pipe in the direction of the e^- beam, the x axis points inward roughly towards the centre of the LEP ring, and the y axis points nearly vertically upward. The polar angle θ is defined as the angle from the $+z$ axis and the azimuthal angle φ as the angle from the $+x$ axis. In the OPAL coordinate system, the xy plane of the detector is referred to as the $r\varphi$ plane, and the rz plane contains the beam axis.

OPAL consists of several detector systems each of which allows different properties of the particles transversing them to be recorded. In order of increasing radial distance from the beam axis, OPAL consists of the following subdetectors: the central tracking system used to determine the direction and momentum of charged particles and to identify certain particles, the time of flight system which identifies some charged particles and rejects cosmic rays, the electromagnetic calorimeter system which measures the energy and direction of photons and electrons, the hadronic

[2] Integrated luminosity is simply the time integral of the luminosity delivered by the accelerator. At LEP, it is usually expressed in inverse picobarns (pb^{-1}) where 1pb^{-1} = 10^{-36} cm^2.

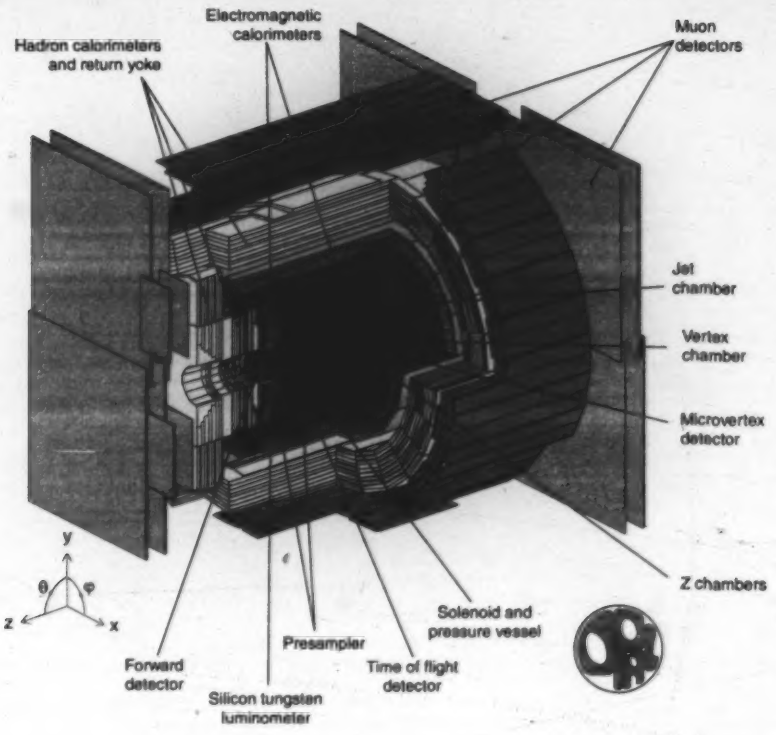


Figure 2.4: A cutaway view of the OPAL detector.

calorimeter system used to measure the energy of hadronic showers, and the muon detector system which identifies muons at the outer edge of the detector. OPAL also includes detectors close to the beam pipe in order to measure the LEP absolute luminosity. A summary of the OPAL subdetectors is given in table 2.1.

2.2.1 The Beam Pipe and Magnet

The OPAL beam pipe is a cylinder of 0.11 mm thickness beryllium at a radius of 53.5 mm from the beam. Beryllium was chosen as the material because of its long radiation length[3]. The OPAL beam pipe contributes only 0.31% of a radiation length

[3]One radiation length, X_0, of a given material is defined as the distance over which a high energy electron travelling in the material loses all but 1/e of its energy due to bremsstrahlung.

Detector System	Subdetector
Central Detector	Silicon Microvertex Detector
	Vertex Chamber
	Jet Chamber
	Z Chambers
Time of Flight System	Time of Flight Barrel Detector
	Time of Flight Endcap Detector
Electromagnetic Calorimeter	Electromagnetic Barrel Calorimeter
	Electromagnetic Endcap Calorimeter
Hadronic Calorimeter	Hadronic Barrel Calorimeter
	Hadronic Endcap Calorimeter
	Hadronic Poletip Calorimeter
Muon Detector	Muon Barrel Detector
	Muon Endcap Detector
Forward Detector System	Forward Detectors
	Silicon Tungsten Luminometer
	MIP Plug

Table 2.1: Summary of the OPAL subdetectors.

of dead material when measured in a direction perpendicular to the beam direction and thus affects the trajectories of particles very little.

The OPAL magnet consists of a solenoidal coil and an iron return yoke. The coil lies at an average diameter of 4.36 m from the beam where the central detector meets the calorimetry and consists of a hollow aluminum conductor fused together with glass-epoxy. The iron return yoke is instrumented for the hadron calorimeter (described in section 2.2.4). The magnet is water cooled and operates at a current of about 7000 A. Within the coil, the magnetic field experienced by the central detector is 0.435 T with a uniformity of ~0.5%. In the calorimetry region between the solenoid and the return yoke, the magnetic field is only a few tens of gauss. This allows the photomultiplier tubes of the electromagnetic calorimetry (described in section 2.2.4) to be operated within their design specifications.

2.2.2 The Central Detector

The OPAL central detector provides all charged particle tracking information for the experiment. The components of the central detector are: the silicon microvertex detector, the vertex chamber, the jet chamber, and the z chambers. These four detectors track charged particles from the interaction point until they exit the magnetic field and enter the calorimetry. With the exception of the silicon microvertex detector, the components of the central detector system are all *drift chambers*. Detectors of this type operate under the principle that a charged particle traversing an active gas volume will ionize the gas and liberate electrons. The liberated electrons then drift in an applied electric field to anode wires and produce a detectable signal which is recorded at a time after the passage of the particle that depends on the characteristic drift time for electrons in the detector gas. If this characteristic drift time is known, the position at which the particle crossed the chamber is calculable from the time taken for the electrons to reach the anode wire.

The vertex chamber, jet chamber, and z chambers all use an active gas mixture consisting of 88.2% argon, 9.8% methane, and 2.0% isobutane. This gas mixture is known within the experiment as "OPAL" gas. The central detector drift chambers are located within the pressure bell and are kept at a pressure of 4 bar. The central detector is also within the 0.435 T field of the magnet solenoid. The curvature in charged particle trajectories due to the magnetic field is used to determine the momentum of the passing particle. Charged particle trajectories from all three drift chambers and the silicon microvertex detector are combined to calculate the final track parameters for each particle. The individual components of the central detector are described below.

The Silicon Microvertex Detector

While not part of the original OPAL detector, the silicon microvertex detector now forms one of the most important parts of the apparatus. Particles with non-negligible lifetimes ($c\tau > 400\,\mu$m) such as B hadrons and τ leptons can travel a few millimeters

before decaying. In order to identify these particles, therefore, it is important for OPAL to be able to resolve the points of decay of these particles, which are separated from the primary e^+e^- vertex of the event. By precisely locating these secondary decay vertices using a detector close to the primary e^+e^- vertex, the lifetime of the original particle may be deduced and used in its identification.

The secondary vertex measurements provided by the silicon microvertex detector have been used within OPAL to produce very precise measurements of the life times of the τ lepton [20] and b flavoured hadrons [21]. As was seen in section 1.2.3, the tagging of b quarks is of particular importance in the search for the Higgs Boson. Thus, the silicon microvertex detector takes an important role in the experimental search presented in this thesis.

As was mentioned in the introduction to the Central Detector system, the silicon microvertex detector is the only detector in the system that is not a drift chamber. Drift chambers are limited in their application to the high-rate environments close to the e^+e^- interaction point as large fluxes of charged particles will saturate the detector gas and prevent useful readout. Solid state detectors, such as the silicon microvertex, do not suffer from this limitation. In a solid state device, the active medium is a solid semiconductor with an applied electric field in order to collect the electron-hole pairs formed by passing charged particles. Semiconductor devices are also much more sensitive to the passage of charged particles since they require on the order of 10 times less energy to produce an electron-hole pair than a drift chamber does to produce an electron-ion pair.

Prior to 1991, OPAL did not contain a microvertex detector, and the vertex drift chamber (see section 2.2.2) was the closest detector to the beam pipe. After the first year of LEP running, the lower than expected background rates near the beam pipe allowed the radius of the pipe to be reduced thus creating space for another subdetector within OPAL. The first version of the silicon microvertex detector was installed in June 1991. This detector was a single sided silicon device and only allowed readout in the φ direction at one fixed radius. The introduction of this detector allowed

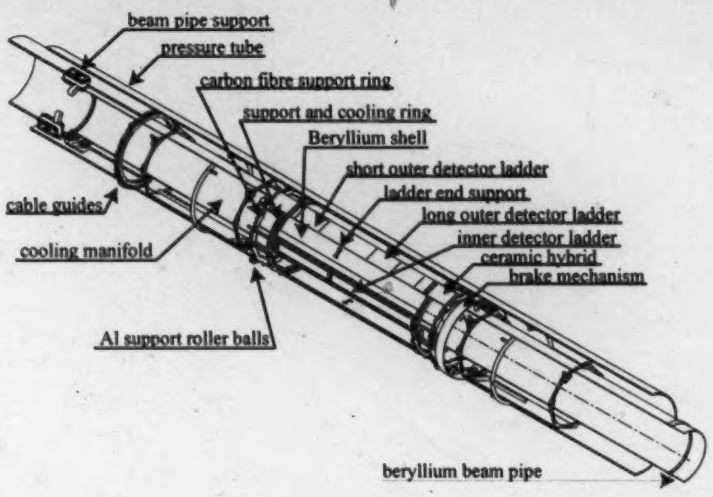


Figure 2.5: Cutaway view of the phase three silicon microvertex detector.

OPAL a single hit position resolution of 8 μm and an impact parameter resolution of 15 μm [22] as compared to an impact parameter[4] resolution of 42 μm prior to the installation of the microvertex detector.

The silicon microvertex detector was converted to a two coordinate readout device in 1993 [23]. This device, capable of readout in both the φ and z directions, was enlarged to cover more of the interaction region in 1995 and 1996. This final version of the detector, dubbed the *phase 3* silicon microvertex detector, was used to take the data used in this thesis.

Figure 2.5 shows a cutaway view of the silicon microvertex detector. The basic detector unit silicon device known as a *wafer*. The wafers are connected end to end into constructions known as ladders which then are arranged in two cylindrical layers to make up the detector. The inner layer consists of twelve 2 wafer "short" ladders at the $+z$ end of the detector and twelve 3 wafer "long" ladders at the $-z$ end of the detector while the outer layer consists of fifteen 2 wafer ladders at the $-z$ end

[4]The impact parameter is a track parameter used to quantify the distance from the primary vertex to a track resulting from a secondary vertex. It is fully described in section 4.4.1.

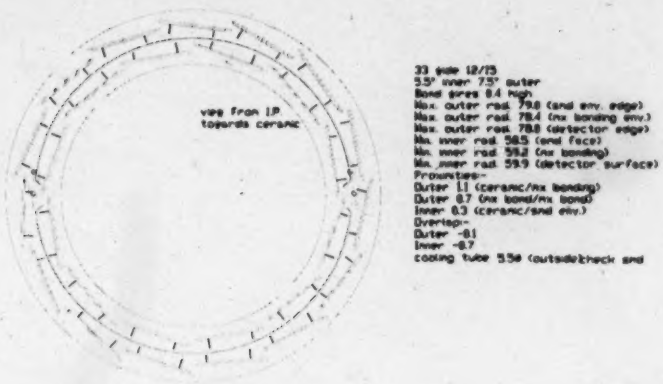

view from IP
towards ceramic
33 wafer 12/15
5.5° inner 7.5° outer
Bond area 0.4 high
Max. outer rad. 79.0 (end env. edge)
Max. outer rad. 78.4 (mx bonding env.)
Max. outer rad. 78.0 (detector edge)
Min. inner rad. 58.5 (end face)
Min. inner rad. 59.2 (mx bonding)
Min. inner rad. 59.9 (detector surface)
Proximities-
Outer 1.1 (ceramic/mx bonding)
Outer 0.7 (mx bond/mx bond)
Inner 0.3 (ceramic/end env.)
Overlap-
Outer -0.1
Inner -0.7
cooling tube 55.0 (outside/check end

Figure 2.6: Cross-sectional view of the silicon microvertex detector showing the overlap in the detector wafers.

and fifteen 3 wafer ladders at the $+z$ end. The ladders are arranged so the centre of the aggregate five wafers is at the e^+e^- interaction point. The inner detector layer is positioned at a radius of 61 mm from the beam while the outer layer is at a radius of 75 mm. The two layers are staggered in φ and the wafers of adjacent ladders are allowed to overlap in order to reduce the gaps due to dead material on the edge of the detector wafers. This overlap and the staggering in φ can be seen in figure 2.6. The single hit efficiency from this arrangement is 97% in the region $|\cos\theta| < 0.89$.

A single "long" ladder, made up of three wafers end to end, is shown in figure 2.7. The detector wafers themselves are made up of back to back single sided silicon strip detectors. They are constructed of high resistivity n-type silicon with n^+ type doping and metalization on the backplane and p type implant strips on the readout side. The side of the wafer facing the interaction region forms the φ detector with the implant strips running in the z direction while the other side forms the z detector where the implant strips run in the φ direction. The p-n-p junction of the detector is formed by running a p type implant drain around the active strips.

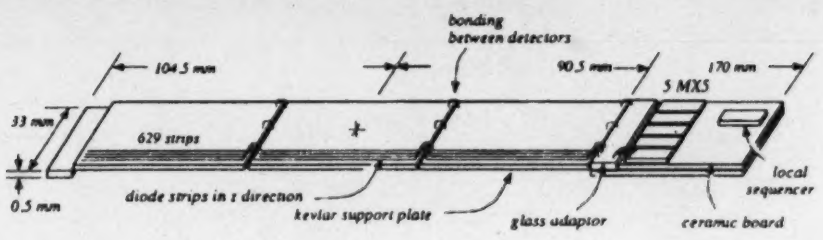


Figure 2.7: Schematic diagram of a "long" three wafer silicon microvertex detector ladder.

Metalized aluminum readout strips are placed above every other φ strip and every fourth z strip and are separated by silicon oxide to form an AC coupling capacitor which allows multiplexing of the detector readout. The detector is multiplexed such that the effective readout pitch is 50 μm in φ and 100 μm in z instead of the physical detector pitch of 25 μm in both φ and z.

A voltage of 35-60 V is applied between the drain line and the backplane. Event though the depletion voltage of the detectors is only in the range of 15-40 V, the extra voltage is necessary in order to compensate for the voltage drop across the strip to drain channel.

The physical support for the detector, also visible in figure 2.5, is made up of three support rings. One of these rings carries the water cooling system for the detector. The detector is sheathed in two semi-cylindrical beryllium shells and fits between the 53.5 mm outer radius of the beam pipe and the 80 mm inner radius of the vertex chamber.

The design of the silicon microvertex detector allows it to present only 1.5% of a radiation length at 90° to the beam direction. This is achieved through the judicious choice of materials used in the detector. Most of the material presented by the detector (> 70%) is comprised of the silicon wafers themselves.

The front end electronics for the φ readout strips are located on 300 μm thick ceramic circuit boards that are located at the end of each ladder. These readout boards are also visible in figure 2.7. The z strip signals are also routed to the

ends of the ladders through a daisy chained gold printed circuit on a borosilicate glass support. This readout arrangement allows all channels to be accessed without presenting unnecessary material to the passing particles.

Online data processing for the silicon microvertex detector is performed using twenty-eight Digital Signal Processors. Pedestals are subtracted channel by channel and then clusters of hits are found. For φ clusters, at least one readout strip must have a pedestal subtracted pulse height at least four times rms noise present in that channel. Strips adjacent to those that pass this test are included in the cluster if their pedestal subtracted heights exceed the rms noise by a factor of 2. The procedure in the z direction is identical, except that two strips on either side of the cluster centre may be included if their pulse heights are two times the rms noise in order to allow for particles traversing the chamber at a shallow angle.

The cluster finding method allows the raw data rate from the silicon microvertex detector to be reduced by a factor of 100. The triggering and online analysis produces a 1.6 ms deadtime which gives a trigger rate of 4-10 Hz. The deadtime of the microvertex detector is very small compared to the overall OPAL deadtime and hence does not cause a readout problem.

The operational temperature of the microvertex detector must be carefully monitored in order to ensure stable operation. Fourteen thermistors gauge the temperature to within 0.005°C. Radiation levels are also monitored by detectors attached to the front end electronics cards. The microvertex is operated in dry nitrogen in order to minimize the contribution to leakage currents due to humidity.

Alignment of the microvertex detector is performed during installation using a precise optical survey and online using $Z^0 \rightarrow \ell^+\ell^-$ events. The lower production rate of these decays at LEP2 energies (above the Z^0 production peak) has made the alignment procedure more reliant on the optical procedure, although the online alignment is still necessary for accurate results. The position uncertainty of the phase 3 microvertex detector within OPAL is estimated to be between 8 and 10 μm in the $r\varphi$ plane and between 10 and 12 μm in the rz plane.

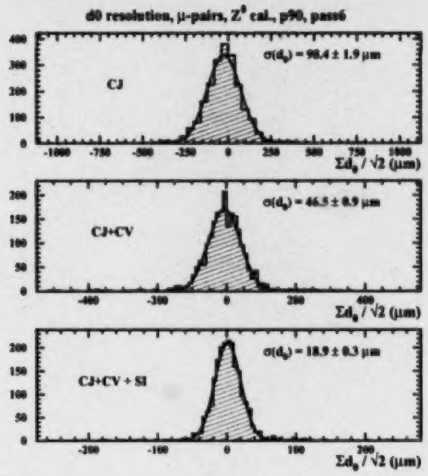


Figure 2.8: Impact parameter resolution for dimuon events collected at the Z peak. The plot shows the resolution with only the jet chamber (top), the jet chamber plus vertex chamber (middle), and all three central detectors in concert (bottom).

The impact parameter resolution of the detector is measured through determining the degree of separation at the interaction point between the two muons from $Z^0 \rightarrow \mu^+\mu^-$ events. The distribution of this measurement around zero gives a measurement of the improvement the silicon microvertex gives to the tracking of OPAL. Figure 2.8 shows the impact parameter distributions in the $r\varphi$ and z directions for various combinations of the silicon microvertex detector (SI), the vertex chamber (CV), and the jet chamber (CJ). The 18.9 μm resolution in the $r\varphi$ direction with the microvertex detector can be compared to a value of 46.5 μm when the silicon microvertex is not used in the impact parameter determination.

The Vertex Chamber

The OPAL Vertex Chamber [24] is a one meter long cylindrical drift chamber that occupies the radial region between 88 and 235 mm from the beam. A schematic drawing of the chamber is shown in figure 2.9. The purpose of the vertex chamber is

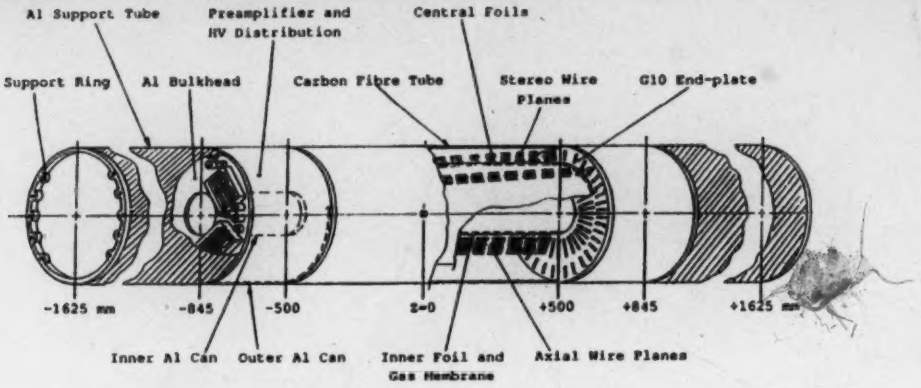


Figure 2.9: Schematic drawing of the OPAL central vertex chamber.

to measure the position of the primary and secondary vertices from particles of non-negligible lifetime and to add to the momentum resolution of the central detector as a whole. Prior to the installation of the silicon microvertex detector, the vertex chamber performed all secondary vertex finding in OPAL.

The wire layout of the vertex chamber is shown in figure 2.10. Between the radii of 90 and 170 mm, the chamber is divided into 36 axial sectors each instrumented with 12 anode wires strung parallel to the beam. The anode wire spacing is 5.3 mm. At radii larger than 170 mm, the chamber is divided into 36 "stereo" sectors with six anode wires each strung 5 mm apart at a stereo angle of 4°. Anode wires in both the normal and stereo sectors are constructed from a 20 μm diameter alloy of tungsten and rhodium. The centre of each readout cell is defined by the row of anode wires while its boundaries are marked by the cathode and potential wires which apply the drift field to the chamber. The potential wires are made from a gold plated copper-beryllium mixture and are 200 μm in diameter while the cathode wires are 125 μm in diameter and are unplated.

Like all drift chambers, the *left-right ambiguity* is a problem in the vertex chamber as the drift time of a passing particle does not specify on which side of the anode wires the particle passed. In order to fix this problem, the anode wires are alternat-

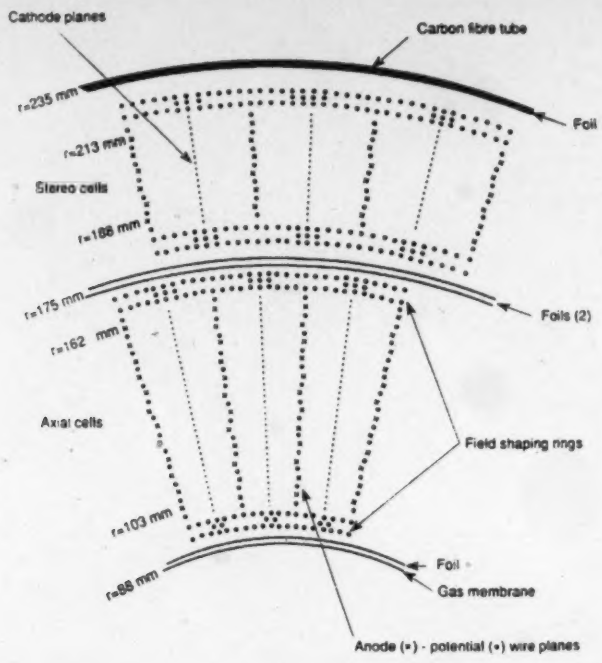


Figure 2.10: Wiring layout of the central vertex detector. The page is in the $r\varphi$ plane.

ingly offset from the cell centering by $\pm 14\,\mu$m so that a set of drift times will only reconstruct a good track on the correct side of the anode wires.

The signals from the anode wires enter preamplifiers at both ends of the wire and are then used to trigger constant fraction discriminators. The average in the arrival time of the signals from each end of the wire is then digitized with a bin size of 0.67 ns. A crude z position measurement is made from the difference in the arrival time of the signals from each end of the anode wire. The timing alone gives a position resolution of $26\,\mu$m in $r\varphi$ and 4 cm in z.

The true performance of the vertex chamber is determined by the detection efficiency and the position resolution. The detection efficiency is affected by the value of the applied anode surface field that determines the gas gain of the chamber. The vertex chamber is operated at an anode surface field of 360 kV/cm which gives a measured single hit efficiency of close to 97%. The optimal drift field for the chamber

is found to be 2.5 kV/cm which corresponds to a drift velocity of 39.5 μm/ns in the "OPAL" gas in the chamber. This drift velocity corresponds to a spatial resolution of 50 μm in $r\varphi$. Loss of spatial resolution due to electron diffusion is kept to a minimum as the maximum drift distance within the chamber is only 1.5 cm. Particles as close together as 2 mm can be resolved by the chamber.

The Jet Chamber

The OPAL jet chamber [25] is the largest part of the central detector system and is responsible for the accurate measurement of particle momentum and particle identification through energy loss per unit distance (dE/dx). This subdetector is also of importance for the Higgs search presented in this thesis as the jet chamber performs much of the work needed for jet finding in OPAL. The jet chamber is a four meter long cylindrical jet chamber with inner and outer radii equal to 0.25 m and 1.85 m respectively. The chamber consists of 24 azimuthal sectors each of which contains a single anode wire plane of 159 wires. The anode wires are spaced at 1 cm intervals and are strung parallel to the beam direction. As in the vertex chamber, the left-right ambiguity is resolved through alternately displacing the anode wires by $\pm 100\,\mu$m from the sector center line. The chamber is instrumented from $r = 25.5$ cm to $r = 183.5$ cm and has a maximum drift distance ranging between 3 cm for the innermost anode wires to 25 cm for the outer wires. The optimal coverage region for the jet chamber is $43^\circ < \theta < 137^\circ$ where up to 159 points can be measured on a single track. At least eight points can be measured per track over 98% of the 4π radians coverage of the detector.

Signals from both ends of the anode wires are collected by preamplifiers mounted on the jet chamber endplates. These signals are then digitized by 100 MHz FADC's[5]. These extremely fast digitizers allow both the drift time accuracy and the two particle separation of the jet chamber to be optimized. As can be seen in figure 2.11, the jet chamber is capable of resolving double hits with high efficiency above a separation

[5]Flash Analog to Digital Converters.

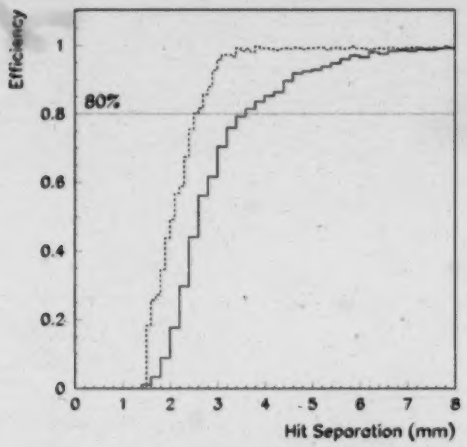


Figure 2.11: Efficiency for double hit detection of closely spaced tracks in the jet chamber. The efficiency is 80% for track spacings above 2.5 mm for hits alone (dashed line) and 80% for spacings above 3.5 mm for hits associated with a track (solid line).

of 2.5 mm. The digitized signals are then processed online for fast track finding to be used in the OPAL trigger. This online processing does a shape analysis on the charge distributions from the FADC's in order to determine the drift time for a given wire hit and hence the $r\varphi$ position. A crude z position measurement is also possible using the ratio of the integrated charge collected on each end of the anode wire to determine the longitudinal position of the hit. The accuracy of this charge division method is dependent on the drift field in the chamber as higher gas gain serves to increase the total amount of charge collected on the anode wires.

The high gas gain required for z resolution is at odds with one of the other major tasks required of the jet chamber. In order to be able to efficiently identify charged particles through the use of energy loss (dE/dx), the chamber must be operated in a drift field low enough to prevent saturation effects where ions formed in the avalanche[6] screen the chamber and prevent some of the collected charge from being

[6]The avalanche is the copious production of electron-ion pairs in the high field near the chamber wires.

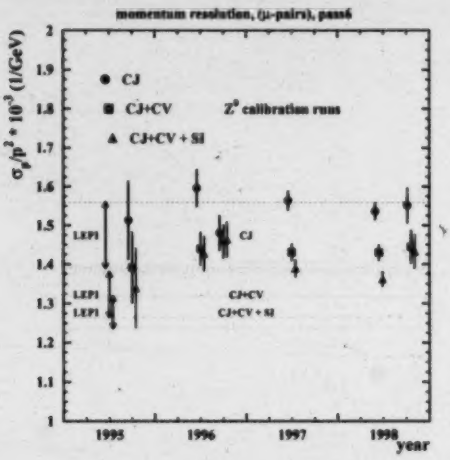


Figure 2.12: The momentum resolution of dimuon events in OPAL versus year of running for the jet chamber alone (CJ), jet chamber plus vertex chamber (CJ+CV), and jet chamber, vertex chamber, and silicon microvertex detector in concert (CJ+CV+SI).

read out. In order to balance these two requirements, the jet chamber is operated at a drift field of 890 V/cm. This field gives a nominal position resolution of 130 μm in $r\varphi$ and 10 cm in z.

This position resolution allows the momentum of particles traversing the jet chamber to be determined with a resolution given by (in the region $|\cos\theta| < 0.7$):

$$(\sigma_p/p)_T^2 = 0.02^2 + 0.0015 p_T^2$$

where p_T is the momentum measured in the xy plane in GeV/c. The limiting factor contributing to the constant term in this equation is multiple scattering in the jet chamber gas. The momentum resolution of OPAL using the jet chamber (CJ), vertex chamber (CV), and silicon microvertex detector (SI) is shown in figure 2.12.

The process of identifying particles through their energy loss per unit distance (dE/dx) is based on the Bethe-Bloch equation [26] for the energy loss of charged particles traversing matter. The dE/dx of a charged particle decreases as $1/\beta^2$ for low values of $\beta\gamma$ to a minimum at $\beta\gamma = 4$ and then rises logarithmically (the *relativistic*

rise) until it saturates at the Fermi plateau[7]. Thus, knowledge of both the dE/dx and the momentum of a given particle can be used to determine its mass and thereby identify it. The efficiency of particle separation in the jet chamber is optimized at a gas pressure between three and four bar and this is one of the main deciding factors in the 4 bar pressure value used in the OPAL pressure bell. Figure 2.13 shows the expected dE/dx behaviour in the OPAL jet chamber for various types of particles as well as actual data points for these particles. A minimum separation in dE/dx of at least 2σ is possible between electrons and pions up to energies of 13 GeV and between pions and kaons or protons up to energies of 20 GeV. The dE/dx resolution for muons pairs is measured to be 2.8% while the resolution for minimum ionizing pions is 3.2% when at least 130 dE/dx measurements are made in the jet chamber.

The Z Chambers

The outermost subdetectors in the central detector system are the Z Chambers [27]. These consist of 24 drift chambers each of which is 4 m long, 50 cm wide, and 59 mm thick. Collectively, the z chambers form a 4 m long barrel surrounding the jet chamber at a diameter of 3.85 m. A cross section of a single z chamber is shown in figure 2.14. The anode wires in each z chamber cell are spaced at 4 mm and are staggered at $\pm 250\,\mu$m from the centre line to resolve the left-right ambiguity. The operational drift field in the chambers is set at 800 V/cm which gives a z coordinate resolution of $\approx 150\,\mu$m. A charge division technique similar to that used in the jet chamber. This technique gives a position resolution in the $r\varphi$ plane of 1.5 cm.

2.2.3 The Time of Flight (TOF) System

In addition to the dE/dx method used in the jet chambers, particles can be identified in OPAL by measuring the time taken for the partices to reach the time of flight system from the interaction point. The Time of Flight (TOF) system, located just outside of the magnet coil within OPAL, is also essential for the OPAL trigger as

[7] β is the velocity of the charged particle in units of the speed of light and $\gamma = 1/\sqrt{1-\beta^2}$.

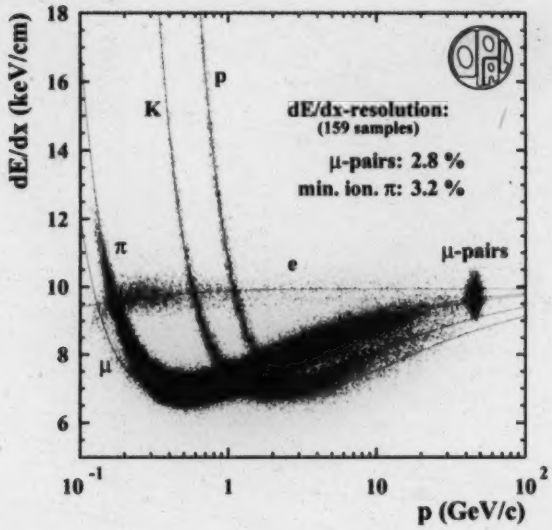


Figure 2.13: Energy loss per unit distance (dE/dx) for various particle types in the OPAL jet chamber. The expected behaviour (lines) and actual measurements (points) are shown for 1994 OPAL data.

it allows the identification of a given event with a unique LEP beam crossing. The TOF system also allows out of time cosmic events to be vetoed by the detector.

The Time of Flight system consists of two subdetectors: the TOF barrel detector (TB) and the Endcap Detector (TE). These two subdetectors are described in the sections that follow.

The Time of Flight Barrel Detector

The Time of Flight Barrel Detector consists of 160 scintillators which form a 6.84 m long cylinder at a radius of 2.36 m from the beam. Each of these scintillators is trapezoidal in cross-section, 6.84 m long, and 4.5 cm thick. The trapezoid has a width of 8.9 cm on the narrow face and 9.1 cm on the wide face. Plexiglass light guides are used at both ends of the scintillators to channel the light to phototubes.

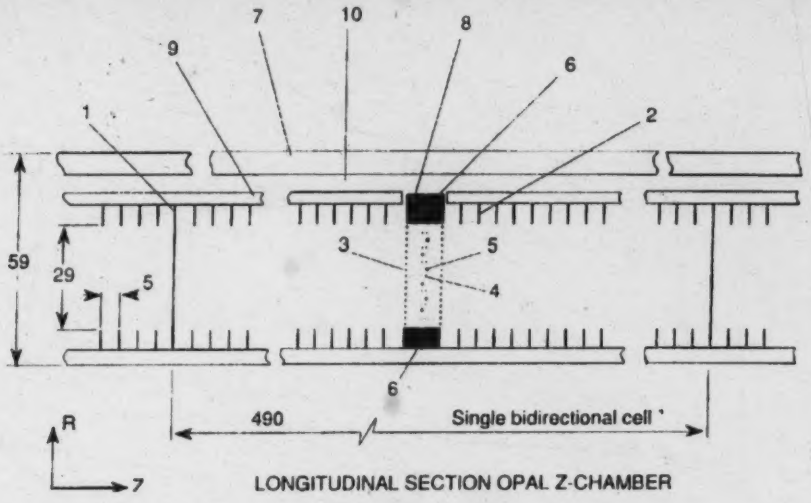


Figure 2.14: A cross-sectional view of a single z chamber.

These tubes can determine both the time of arrival of the signal, and the total charge deposited in the scintillator by the passing particle.

Measurements of the timing resolution of the TOF barrel detector have been performed using dimuon events in OPAL. A timing resolution between 300 and 400 ps results from these tests depending on the z position at which the particle passes.

The Time of Flight Endcap Detector

The time of flight endcap detector [28] was designed as an improvement for LEP2 running in order to provide increased hermiticity and timing resolution for the OPAL detector. The increased bunch collision frequency at LEP2, in addition to the increased importance of new physics containing unbalanced visible momenta, isolated photons, or long-lived particles made the addition of a TOF endcap to OPAL an attractive proposition. The University of Alberta played a major role in the design, construction, and installation of this subdetector.

The Time of Flight Endcap detector is shown in its position in OPAL in fig-

ure 2.15. The detector is made up of two layers of scintillating tiles placed in between the endcap presampler and the electromagnetic calorimeter (see section 2.2.4). This space was left as tolerance in the original OPAL design. The tiles are separated into three annular regions with the innermost region being divided into 24 sectors in φ and the outer two regions into 48 φ sections. This geometry was selected in order to integrate the subdetector into the OPAL trigger matrix while physically allowing it to fit over the electromagnetic endcap calorimeter. One of these 1/24 sectors is also visible in figure 2.15.

The tiles themselves are constructed from 10 mm thick Bicron BC408 scintillator and are embedded with 1 mm diameter wavelength shifting (WLS) fiber. The fiber has an emission peak at 500 nm and is embedded in a geometry designed to maximize the light collection. The WLS fibers connect to clear fiber that transmits the collected light about 15 m to the phototubes.

The Time of Flight Endcap detector has been successfully used to determine the collision time, t_0, in OPAL. The LEP bunch spacing is measured to be 335 ± 1 ns as compared to an expected value of 334 ns, and the time jitter of the subdetector is 3.0 ns as compared to the design value of 5.0 ns.

2.2.4 The Calorimeter System

The OPAL calorimeter system is designed to accurately measure the energies of electrons, photons, and hadrons using a variety of devices. The calorimetry is also important to the analysis presented in this thesis since it is used in jet finding and constrained kinematic fitting techniques. The calorimeter system is broadly divided into the electromagnetic system and the hadronic system. Both are described in the sections that follow.

The Electromagnetic Calorimeter

The electromagnetic calorimeter is capable of measuring the energies of electrons and photons ranging from tens of MeV to over 100 GeV. The calorimeter covers

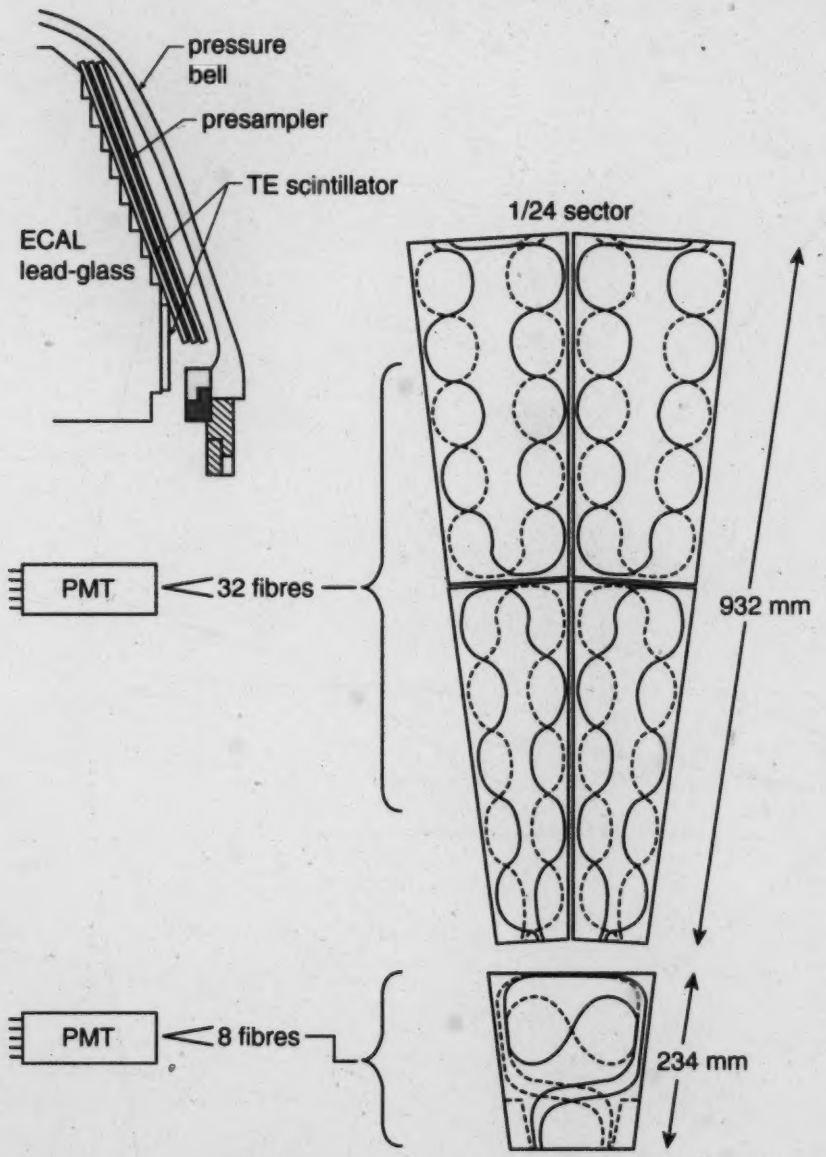


Figure 2.15: The Time of Flight Endcap detector. The position of the subdetector in OPAL is shown as well as a schematic of a single detector sector.

over 98% of the solid angle of OPAL and is divided into two sections: the barrel detector which covers $|\cos\theta| < 0.82$ and two endcap regions that cover the region $0.81 < |\cos\theta| < 0.98$. Both the barrel and the endcap electromagnetic calorimeters make use of a system of preshower counters and a main calorimeter constructed of lead glass blocks. The preshower detectors are necessary to accurately determine the energy loss of showers that begin before the electromagnetic calorimeter.

The barrel electromagnetic presampler [29] consists of sixteen streamer chambers arranged in a 6.623 m long cylinder at a radius of 2.388 m from the beam. Each of the chambers contains two layers of streamer tubes each of which has a square cross section with area 9.6 mm^2 and is instrumented with 1 cm wide cathode strips on both sides of each tube layer. The cathode strips are inclined at an angle of 45° to the wire direction and are orthogonal to each other in the inner and outer layers of tubes. A radial position resolution between 4 and 6 cm, depending on the incident particle energy, is achieved with the presampler while a rough 10 cm z position resolution is made possible through a charge division technique.

The main section of the barrel electromagnetic calorimeter consists of 9440 lead glass blocks each of which is covered in vinyl fluoride and measures 10 cm×10 cm×37 cm. The blocks are arranged just outside of the TOF detector with an inner radius of 2.455 m from the beam. The blocks point towards the interaction region in order to limit the losses due to block interfaces, and each presents about 24.6 radiation lengths to particles passing radially. Readout of the lead glass blocks is achieved using magnetic field tolerant phototubes.

The intrinsic energy resolution of the lead glass blocks, in the absence of extra material, is given by:

$$\frac{\sigma_E}{E} = 0.2\% + \frac{6.3\%}{\sqrt{E}}.$$

In OPAL, however, this optimal resolution is reduced due to the extra material preceding the electromagnetic calorimeter. The measured energy resolution of the barrel electromagnetic calorimeter is $\sigma_E/E = (1.2 \pm 0.3)\% + (20.8 \pm 1.8)\%/\sqrt{E}$. A

rough position measurement is also made possible due to the granularity inherent in the lead glass blocks. This resolution is measured to be 5 mm in the barrel region.

The endcap electromagnetic presampler is comprised of 32 thin multiwire proportional chambers (MWPC) that are placed between the pressure bell and the electromagnetic endcap lead glass calorimeter. They cover the annular region between $0.83 < |\cos\theta| < 0.95$ and have a total thickness of only 6.8 mm. The angular position resolution afforded by the endcap presampler is measured to be 4.6 mrad.

The endcap lead glass calorimeters are similar in design to their barrel counterpart. Each endcap is umbrella shaped and consists of 1132 lead glass blocks. In order to follow the contour of the pressure bell, the endcap blocks come in three sizes each of which presents around 22 radiation lengths to incoming particles. The lead glass blocks are read out through single stange magnetic field tolerant phototubes that are attached directly to the blocks. This is very important in the endcap, as some of the phototubes operate within the full magnetic field of the solenoid. The electromagnetic endcap calorimeter has a measured energy resolution of $(4.6 \pm 0.7)\% + (12.5 \pm 3.6)\%/\sqrt{E}$.

The Hadronic Calorimeter

The OPAL hadronic calorimeter is used to measure the energy of hadrons that pass through the electromagnetic calorimetry as well as to complement the muon chambers in the identification of muons. The calorimeter itself is instrumented within the iron return yoke of the OPAL magnet, and is divided into three sections: the barrel, the endcaps, and the poletips. The iron used in the hadronic calorimeter presents at least four interaction lengths to incident particles over 97% of the solid angle of the detector.

The barrel hadronic calorimeter surrounds the electromagnetic calorimeter at radii between 3.39 and 4.39 m and is 10 m long, thus covering the polar angles $|\cos\theta| < 0.81$. It consists of eight 10 cm thick iron slabs interspersed with nine 2.5 cm thick layers of active streamer tubes. In order to increase the number of

radiation lengths presented by the barrel, the central 4 m of the detector is covered by an additional two iron plates.

The anode wires in the streamer tubes run parallel to the beam axis in an environment of 75% isobutane and 25% argon. Ten millimetre wide cathode strips are placed on one side of the tubes, while 4 mm wide aluminum strips run parallel to the anode wires on the other side of the tube. The readout combines hits on clusters of cathode pads to form calorimeter *towers*. The aluminum strips give a φ resolution of 2.5 mrad.

The hadronic endcap calorimeter covers the region $0.81 < |\cos\theta| < 0.91$ and consists of seven 10 cm thick irons slabs separated by 3.5 cm wide instrumented gaps. The gaps in the hadronic endcap calorimeter are also instrumented with streamer tubes with a similar configuration to those found in the barrel except that the aluminum strips are arranged perpendicular to the beam.

A total of 976 calorimeter towers are present when the readout of the hadronic barrel calorimeter cathode pads are combined with those of the endcap. These towers divide the solid angle of OPAL into 48 bins in φ and 21 bins in θ and are used for fast trigger development. The strip readout, on the other hand, is used to map shower profiles and for muon tracking.

The pole tip calorimeters [30] were installed to extend the reach of the hadronic calorimeter between $0.91 < |\cos\theta| < 0.99$. Each pole tip is constructed from nine 8 cm thick irons slabs separated by gaps of 1 cm. The gaps contain thin multiwire proportional chambers (MWPC's) that use a mixture of 55% carbon dioxide and 45% n-pentane as the gas mixture. The chambers also have cathode pads on one side and strips on the other. The cathode pad area is $500\,\mathrm{cm}^2$ while the strips are 2.5 cm wide.

The number of radiation lengths present in the hadronic calorimeter guarantees that no more than 0.1% of pions will pass through the detector into the muon chambers. In fact, only 0.8% of all particles at 5 GeV/c will punch through the hadronic calorimeter. The average energy resolution of the hadronic calorimeter is

$120\%/\sqrt{E}$.

2.2.5 The Muon System

The OPAL muon detector is responsible for the identification of muons in the presence of hadronic background. It is constructed as large-area drift chambers and is divided into a barrel and two endcaps. At least one layer of the muon chambers covers 93% of the OPAL solid angle.

The barrel portion of the muon chambers consists of four layers of 10 m long by 1.2 m wide drift chambers. This forms a cylinder around OPAL at a radius of 5 m. At least one layer of the chambers covers the region $|\cos\theta| < 0.68$ while all four layers cover the region $|\cos\theta| < 0.72$. There are a total of 110 drift chambers in the barrel portion of the subdetector each of which contains two drift cells each with one anode wire running parallel to the beam axis. Each drift cell is 0.6 m wide, 1.5 cm thick, and about 10.4 m in length. The drift time is used to calculate the $r\varphi$ position of a passing track to a resolution of 1.5 mm while diamond shaped cathode pads give the z coordinate to within 2 mm.

The muon chamber endcaps cover the polar region $0.67 < |\cos\theta| < 0.98$ and consist of 12 m×12 m panels placed perpendicularly to the beam axis. Each panel is instrumented with four layers of streamer tubes using 10 mm strip readouts and a gas mixture of 75% isobutane and 25% argon. The resolution of these chambers is similar to that of the muon barrel at 1 mm in both the x and y coordinates.

The identification of muons in OPAL is performed by extrapolating central detector tracks to the interface with the muon chambers where they are matched with track segments from the chambers of the muon system. A correction for multiple scattering in the detector is applied before the track matching is performed. For muons of interest, this multiple scattering is about 2 mm which indicates that the position resolution of the muon chambers is sufficient to match tracks successfully.

2.2.6 The Forward Detector System

A system of several subdetectors, all placed close to the OPAL beam pipe, are used to measure the absolute luminosity delivered by LEP to the detector. This measurement is of extreme importance, as it allows the accurate calculation of production rates of observed particles and exclusion limits for particle searches. The forward detectors measure the LEP luminosity through the detection of low angle Bhabbha scattering events. Three sub-systems make up the forward detector system: the forward detectors, the silicon tungsten luminometer, and the MIP plug.

The Forward Detectors

The forward detectors consist of a forward calorimeter, the tube chambers, the gamma catcher, and the far forward monitors. The forward calorimeters surround the OPAL beam pipe at each of the detector by filling the holes left at the centre of the electromagnetic endcap calorimeters. Each consists of 35 sampling layers of lead-scintillator sandwich and is divided into 16 azimuthal segments. The calorimeter has an integrated preshower that presents 4 radiation lengths, while the calorimeter proper presents 20 radiation lengths. The calorimeter is capable of a position measurement using its granularity to give a polar angle resolution of between 4 and 10 mrad. The energy resolution of the forward calorimeter has been measured to be $18\%/\sqrt{E}$.

In between the forward calorimeter presampler and the lead-scintillator are three layers of brass proportional tubes. These give individual shower positions to an accuracy of ± 2 mrad and are used to augment the cluster information given by the main calorimeter.

Two small electromagnetic calorimeters, known as the gamma catchers, are used to cover the region in θ from 143 to 193 mrad between the forward calorimeter and the inner interface of the electromagnetic endcap calorimeter. These units consist of eight azimuthal segments each of which is made from lead-scintillator sandwich. The gamma catchers only present seven radiation lengths of material and hence must

be used in conjunction with either the forward calorimeters or the electromagnetic endcap to fully contain showers.

Mounted at 7.85 m from the interaction point along the beam pipe are small lead-scintillator sandwich detectors mounted on either side of the beam pipe. These detectors measure 50 mm×150 mm and are 20 radiation lengths thick. The main function of these far forward monitors is to keep watch over beam related backgrounds.

The silicon tungsten calorimeter was a more recent addition to OPAL and was commissioned in 1993. It consists of two cylindrical calorimeters which fit around the beam pipe at ±238.94 cm from the interaction point. Each detector is made up of 19 layers of active silicon separated by 18 layers of tungsten and uses a bare layer of silicon at the end facing the interaction point as a preshower. The energy resolution of the silicon tungsten calorimeter has been measured to be $28\%/\sqrt{E}$.

The Minimum Ionizing Particle (MIP) plug was a recent addition to the OPAL detector and was installed at the same time as the time of flight endcap subdetector. Tiles similar to those used in the time of flight endcap are used in the MIP plug with each endcap containing 32 10 mm thick tiles. The tiles are arranged as an inner and outer pair of annular layers with each layer divided into 8 sectors. A schematic of the MIP plug tile layout may be seen in figure 2.16. The main purpose of the MIP plug is trigger and veto functions in the forward region.

2.2.7 Triggering and Online Data Acquisition

Interesting physics processes are separated from backgrounds by the OPAL trigger system. This allows the data rate collected by OPAL to be reduced dramatically, and hence reduces the difficulty of data storage and handling.

Triggers in OPAL are one of two types: stand alone triggers such as multiplicity counts and energy sums, and threshold signals from the $\theta\varphi$ binning of the trigger matrix. The trigger matrix receives inputs from the central trackers, the time of flight counters, the electromagnetic calorimeter, and the muon detector. A list of

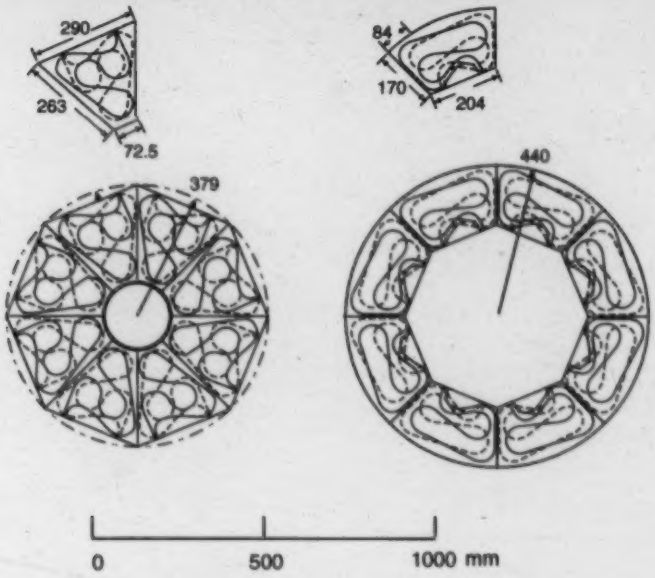


Figure 2.16: The tile layout of the minimum ionizing particle (MIP) plug.

the correlations between trigger matrix elements that are sufficient to cause a trigger are listed in the *trigger definition file* that can be updated online for any OPAL run.

The central trigger unit contains a dedicated trigger bus for fast uninterrupted operation. When a trigger is tripped within OPAL, the General Trigger Unit (GTU) sends signals to the Local Triggers Units (LTU) on each subdetector which in tern read out their data while the LTU is held in a busy state. Only when all the LTU's have reported to the GTU that readout is completed does the GTU resume its ready state. If no trigger is indicated for a given bunch crossing, a rest signal is sent to each of the subdetectors. The OPAL trigger reduces the data rate to about 10 Hz.

After a successful trigger, each of the OPAL subdetectors is read out into its Local System Crate (LSC). This data is then assembled in the Event Builder (EVB) which is connected to each of the LSC's by high speed memory map links. After it is assembled in the EVB, the data is filtered in order to compress the data and reject badly measured events due to subdetector failure. Header information for all

triggers, however, is kept on disk regardless of whether or not they are filtered at this stage.

Good events that pass the filter are written to disk in 20 MB partitions and are also transferred to the online reconstruction farm where calibration constants are applied[8]. Reconstructed events are backed up on optical disk and tape cartridge and are typically available for offline analysis within hours of the data being taken.

[8]The OPAL reconstruction program ROPE is used for this analysis.

Chapter 3

Artificial Neural Networks

Separating a signal from various background sources in particle physics usually involves the selection of several discriminating variables, and then the application of successive cuts on these variables in order to optimally select the signal while rejecting the background. For more than a few variables, selecting the optimal cuts can be a very time consuming task, and taking correlations between the variables into account is very difficult. An Artificial Neural Network (ANN), on the other hand, allows discriminating variables and their correlations to be considered together in order to produce a likelihood that a given event belongs to either the class of signal events, or the class of background events.

ANN's are particularly well suited to particle physics analysis as they are capable of recognizing multi-dimensional features in the input data that might be missed with a conventional multivariate analysis. The development time for an ANN analysis is usually shorter than a comparable conventional analysis due to the fact that ANN's are trainable, that is they can self-organize and optimize their selection capability when presented with known classes of data. ANN's are also fault tolerant in that noisy or incomplete data will not always destroy the power of the analysis as the ANN can draw upon other features in the data to make a decision. ANN's have been implemented very successfully in many areas of particle physics [31]. Some of these are described in section 3.7.

An Artificial Neural Network [32] consists of a matrix of interconnected units

called *neurons*. A neuron is the basic computation unit of the neural network and combines the inputs from several other neurons in determining its output state. Discriminating variables for a given analysis are presented to a layer of input neurons which process and pass on various features of the data to other layers. ANN's can consists of several intermediate layers of neurons before a final classification decision is made.

There are some drawbacks to the use of ANN's in particle physics analysis. Since the decision making process happens in a "black box" with no visible intermediate steps, the actual relationships between the input variables which the ANN uses to make decisions are hidden from the user. Thus, systematic error checking using ANN's can be difficult. Also, great care must be taken in the training of ANN's in order to guard against non-representative features being learned by the network which will skew the results of the analysis. Despite these drawbacks, well designed ANN's can be very beneficial to the analysis efforts of the particle physicist.

3.1 The Neuron

All Artificial Neural Networks are based upon collections of *neurons*. The artificial neuron can be implemented in either a hardware (FPGA[1]) or software format. The type of neuron most used in particle physics applications is the so-called *dot product* neuron that is pictured in figure 3.1. The state of such a neuron is completely described by its activation function f and its current state s_i. When neurons are connected together to form an ANN, the connections are assigned weights with the weight between neurons i and j being given by w_{ij}.

The state of a dot product neuron is determined by performing a dot product P_i of the states of the neurons connected to it with their associated weights.

$$P_i = \sum_j w_{ij} \cdot s_j$$

[1]Field Programmable Gate Array.

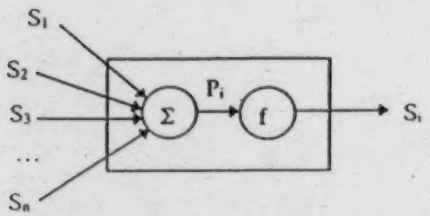


Figure 3.1: The dot product neuron. The neuron calculates the dot product of the states of the incoming neurons s_i with their respective weights w_{ij} and then uses this value to calculate the output of the neuron through the activation function f.

This product is then passed through the activation function f in order to determine the output state of the neuron.

$$s_i = f(P_i) = f(\sum_j w_{ij} \cdot s_j) \tag{3.1}$$

The activation function f is chosen from several common functions and consists of applying a threshold of some sort, and then some amplification to the dot product P_i. In early ANN's, the most common activation function was a simple step function that mimicked the on/off behaviour of digital systems. In contemporary networks, however, differentiable functions such as the logistic function

$$\sigma(z_i) = \frac{1}{1 + e^{-z_i}} \tag{3.2}$$

are used as the activation function. This sort of activation function helps the classification process by dividing the feature space into a region where the activation function is approaching zero, and a region where it is approaching one. Thus, cuts in the multi-dimensional feature space of the network can be made by the ANN

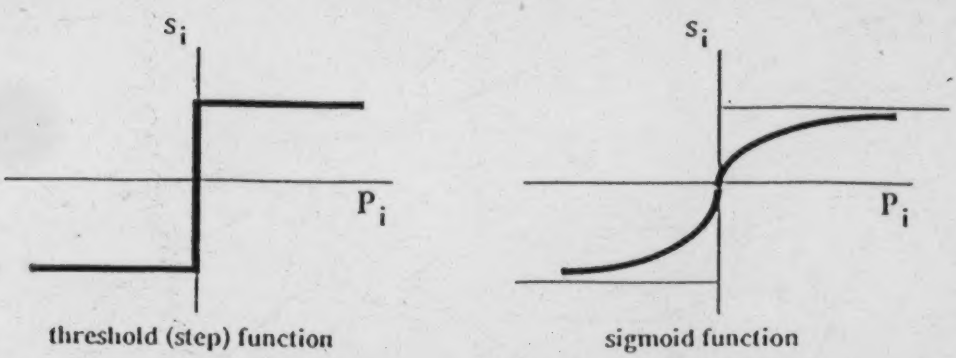


Figure 3.2: The step activation function (left) and the nonlinear sigmoid function (right).

with this activation function. Both the step function and the sigmoid are shown in figure 3.2.

A second type of neuron, the distance neuron, is sometimes used in Artificial Neural Networks. Instead of calculating the weighted sum of the input neuron states, the "distance" between the input weight vector $\vec{W}$ and the input state vector $\vec{S}$ is calculated. This is commonly calculated through the euclidean norm.

$$d_i = ||\vec{W} - \vec{S}||$$

For this type of neuron, the activation function f is a bell curve and hence neurons for which the input weights match the input states more closely will have the highest activation. This type of neuron is used most commonly in competitive learning methods and unsupervised self-organizing networks [33]. Since this type of network is not usually used in particle physics applications, the distance neuron will not be further considered in this thesis.

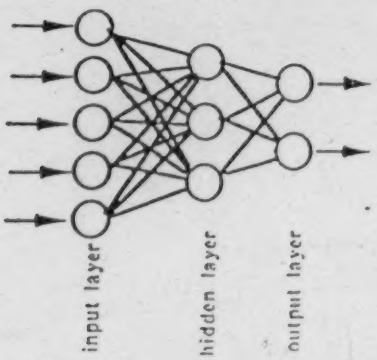


Figure 3.3: The architecture of a feed forward network showing the input, hidden, and output layers.

3.2 Structure of an ANN

In building an ANN, attention must be paid to the number of neurons it contains and the particular way in which they are connected. There are two main architectures of ANN: the *feed-forward network* and the *recurrent* network. In the feed-forward model, neurons are arranged in a series of well-defined layers where each neuron belongs to only one layer. Each neuron in a given layer is connected to each neuron in adjacent layers, but there are no connections between neurons in non-adjacent layers of the network. This arrangement is shown in figure 3.3. As can be seen in the figure, the feed-forward network consists of well defined input and output layers with a number of hidden layers in between. At the input layer, raw input data is accepted into the network, and the results of processing through the hidden layers are presented at the output layer. This type of network is also known as a *Multi-Layer Perceptron* (MLP) and is the type that is used most often for classification tasks in particle physics. It is networks of this type that will concern us in this thesis.

Recurrent networks, on the other hand, feature connections between each and every neuron in the network and thus have no well-defined layer structure. A recur-

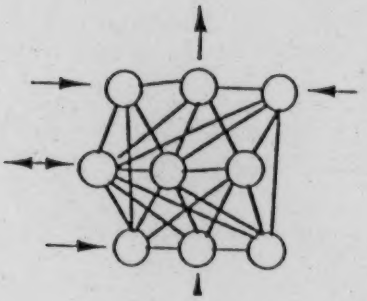


Figure 3.4: The architecture of a recurrent network. Every neuron has a connection to every other neuron in the network.

rent network is shown schematically in figure 3.4. Any neuron in a recurrent network may be an input neuron, an output neuron, or both. Networks of this type are used widely in pattern recognition tasks, and have found some applications in charged particle tracking algorithms in particle physics. Recurrent networks are useful for extracting pattern information even in the presence of noise or missing data, and hence are well suited to track finding in particle physics experiments.

3.3 Significance of Hidden Neurons

The way in which Artifical Neural Networks can create cuts in the multi-dimensional feature space of input data can be shown with two examples. In general, a feed-forward ANN can create "feature surfaces" in $n + 1$ dimensions where n is the number of hidden layers in the network.

In the first example, two classes of data $\mathcal{A}$ and $\mathcal{B}$ are distinguishable due to the fact that they lie on opposite sides of the line $aX + bY + c = 0$. A two layer ANN, shown in figure 3.5, is applied to this situation that uses the values of X and Y as inputs and a threshold input of value c to make a total of three input nodes. These

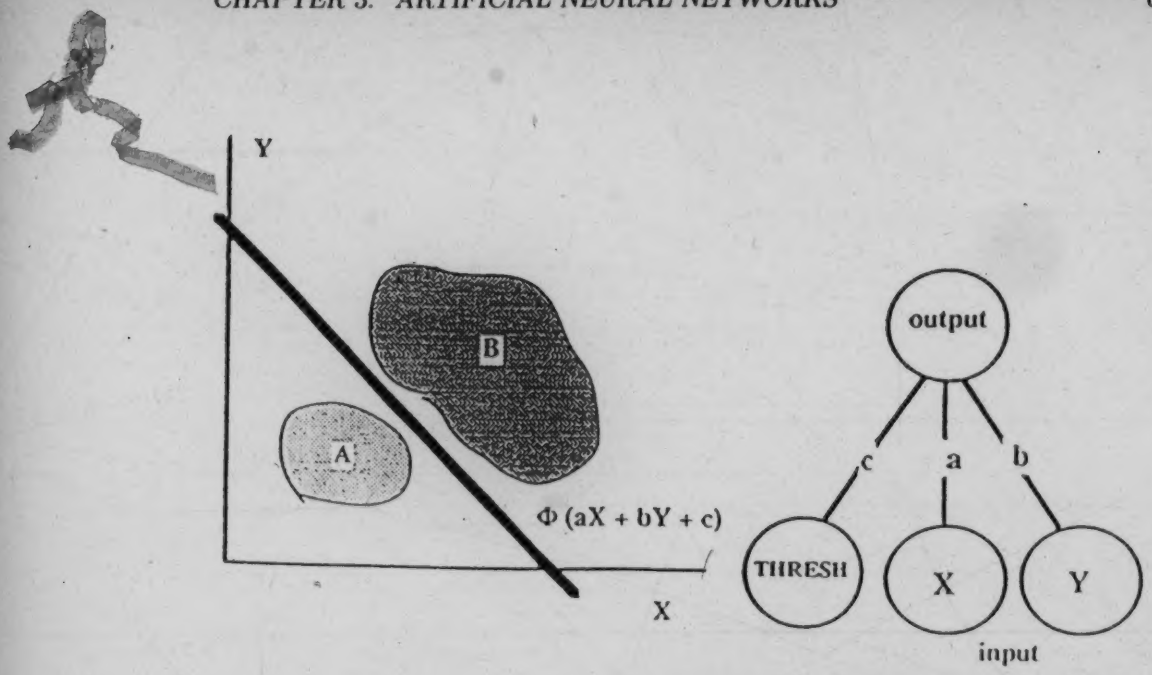
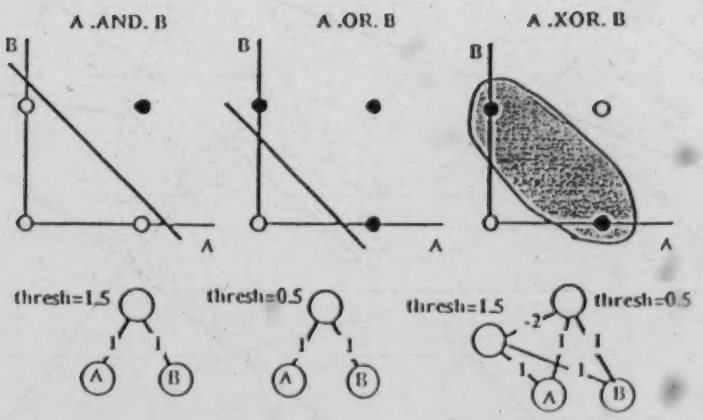


Figure 3.5: A simple example of a two layer ANN used to separate two classes of data $\mathcal{A}$ and $\mathcal{B}$ (left). With no hidden layer, the ANN is capable of only a one dimensional cut in feature space along the line $aX + bY + c = 0$. The network architecture is shown on the right. The threshold input node always reports a value of c to the network.

nodes then feed into a single output node that uses a simple step activation function of the type $\Phi(aX + bY + c)$. This activation function allows the two data sets to completely separated by the network.

In the second example, shown in figure 3.6, an ANN is designed to create a two dimensional decision surface to emulate the exclusive or (XOR) function. As can be seen in the figure, no single line can separate the data sets A.XOR.B=true and A.XOR.B=false from each other. By the introduction of a single hidden layer that determines whether inputs A and B are *both* true, a two dimensional decision surface is created by the network that allows unambiguous identification of the class of data. It is easy to see how more complex networks allow relationships between the input data to be exploited for classification purposes.



A	B	Hidden Neuron	Output
0	0	0	0
0	1	0	1
1	0	0	1
1	1	1	0

Figure 3.6: The two dimensional space containing the outputs of A.XOR.B is shown at the top of the diagram. Points for which A.XOR.B is true are shown as dark circles, while points for which A.XOR.B is false are shown as open circles. In the top left and top middle plots, it is shown that a one-dimensional decision surface is insufficient to separate the data classes. The addition of a hidden layer, however, that determines the state of A and B together, makes classification possible. The chart at the bottom of the figure shows the neuron outputs for various inputs A and B.

3.4 The Learning Process

Feed-forward networks used in data classification tasks are usually *supervised* networks in that they are trained to recognize a certain class of data. This training is achieved using simulated data of which the class is known and that matches the properties of the unknown data very well.

The network being trained is presented with a target output vector $\vec{T}$ that corresponds with the class of data being given to the ANN. For example, in the simplest case of a single output neuron, the output target would be 1 for signal events and 0 for background events. A set of training data is then fed to the network and the output of each event is compared to $\vec{T}$. The weights between the neurons in the ANN are then updated to allow the output vector $\vec{O}$ to more closely match the target $\vec{T}$. This change of the network weights, Δw_{ij}, with time is governed by several principles.

The first of these principles, known as *Hebbian Learning* embodies the notion that the weight between two given neurons strengthens in proportion to the product of the states of the two connected neurons. This strengthening can be described by the equation:

$$\Delta w_{ij} = \eta s_{it} s_{jt} \tag{3.3}$$

where s_{it} and s_{jt} are the states of the two neurons in question and η is called the "learning parameter". Hebbian learning ensures that only connections that are successfully identifying data features continue to develop weights.

The second principle, *Delta Rule Learning*, gives the mechanism whereby the ANN attempts to reduce the differences between the output vector $\vec{O}$ and the target vector $\vec{T}$ over the training events. For a given training event p, a quadratic form E_p can be calculated to quantify the degree of difference between $\vec{O}$ and $\vec{T}$.

$$E_p = \frac{1}{2}\left|\vec{O}_p - \vec{T}_p\right|^2 = \frac{1}{2}\sum\left[O_p(i) - T_p(i)\right]^2 \tag{3.4}$$

The global error function E is obtained by summing E_p over several training events.

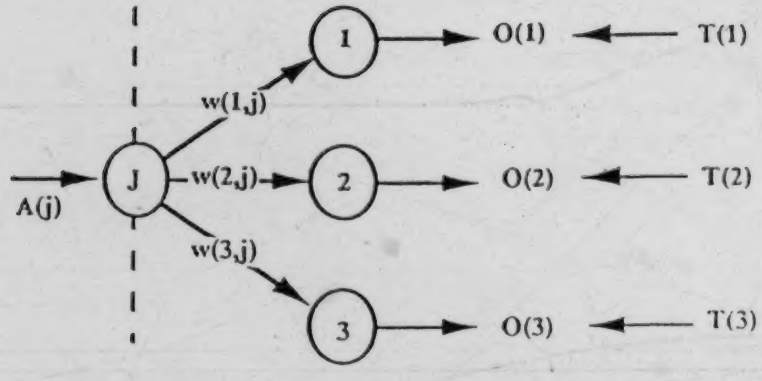


Figure 3.7: A simple two layer ANN.

The Delta Rule then states that this global error function should be minimized and that changes the weights w_{ij} in proportion to the effect they have on the minimization process. This evolution of the weights is known as *gradient descent* and is described symbolically by the equation

$$\Delta w_{ij} = -\eta \frac{\partial E}{\partial w_{ij}} \tag{3.5}$$

where η is again a constant called the learning parameter.

In order to clarify these ideas, we now consider the simple two layer ANN shown in figure 3.7. This ANN consists of a single input node which receives input vectors $\vec{A}_p$ and three output nodes with states $O(1), O(2)$, and $O(3)$ respectively. Weights between the input node and the output nodes are given by w_{ij} and a target output vector $\vec{T}_p = (T(1), T(2), T(3))$ is known for each of the p training events. All neurons in the example are of the dot product type.

Recalling the output of a dot product neuron in equation 3.1, we can calculate

the components of $\vec{O}$ for our simple ANN using the equation:

$$O_p(i) = f\left[\sum w_{ij} A_p(j)\right] = f\left[s_p(i)\right] \tag{3.6}$$

where f is the activation function of the neuron. Using equation 3.4, we may now calculate the evolution of the error function as follows:

$$\begin{aligned}\frac{\partial E}{\partial w_{ij}} &= \sum_p \frac{\partial E_p}{\partial w_{ij}} \\ &= \sum_p \frac{\partial E}{\partial \vec{O}_p} \cdot \frac{\partial \vec{O}}{\partial s_p(i)} \cdot \frac{\partial s_p(i)}{\partial w_{ij}} \\ &= \sum_p \left[\vec{O}_p(i) - \vec{T}_p(i)\right] \cdot f'\left[s_p(i)\right] \cdot A_p(j). \end{aligned} \tag{3.7}$$

Now, applying equation 3.5, we may write the change in the weighting constants w_{ij} as:

$$\begin{aligned}\Delta w_{ij} &= -\eta \sum_p \left[\vec{O}_p(i) - \vec{T}_p(i)\right] f'\left[s_p(i)\right] A_p(j) \\ &= -\eta \sum_p \delta_p(i) A_p(i) \end{aligned} \tag{3.8}$$

where

$$\delta_p(i) = \left[\vec{O}_p(i) - \vec{T}_p(i)\right] f'\left[s_p(i)\right]. \tag{3.9}$$

Equation 3.8 is known as the *Widrow-Hoff* formula. We can see the effect of Hebbian learning in equation 3.8 as the magnitude of the weight change depends both on the strength of the inputs $A_p(i)$ and the input weights w_{ij} through their contribution to $s_p(i)$. It is also noteworthy that the activation function f must be differentiable in order to apply the Widrow-Hoff formula. This means that continuous functions such as the sigmoid of equation 3.2 must be used.

3.5 Backpropagation

The learning process used in an ANN of more than two layers becomes more complicated as the evolution of the weights must somehow be communicated from the output layer back through the hidden layers. This communication method is called

backpropagation. If we consider the three layer network shown in figure 3.8, we can apply equation 3.5 to see the evolution of the weights w_{ij}.

$$\Delta w_{ij} = -\eta \sum_p \frac{\partial E_p}{\partial w_{ij}} \tag{3.10}$$

and from equation 3.7, we have

$$\frac{\partial E_p}{\partial w_{ij}} = \frac{\partial E}{\partial \vec{O}_p(i)} \cdot f'\left[s_p(i)\right] \cdot A_p(j) \tag{3.11}$$

in exact symmetry with the two layer case. In the case of a three layer network, however, we cannot directly relate the target vector $\vec{T}$ and the output vector $\vec{O}$ since the input and output layers are separated by at least one hidden layer of neurons. Instead, we can calculate $\frac{\partial E_p}{\partial \vec{O}(i)}$ using an expansion:

$$\begin{aligned} \frac{\partial E_p}{\partial \vec{O}_p(i)} &= \sum_k \frac{\partial E_p}{\partial \vec{O}_p(k)} \cdot \frac{\partial \vec{O}_p(k)}{\partial s_p(k)} \cdot \frac{\partial s_p(k)}{\partial \vec{O}_p(i)} \\ &= \sum_k \delta_p(k) w_{ki} \end{aligned} \tag{3.12}$$

where we have recalled the definition of δ_p from equation 3.9.

Now, substituting equation 3.12 into equations 3.10 and 3.11, we obtain:

$$\Delta w_{ij} = -\eta \sum_p \left(f'\left[s_p(i)\right] \cdot A_p(j) \cdot \left[\sum_k \delta_p(k) w_{ki} \right] \right) \tag{3.13}$$

for the evolution of the weights in the hidden layer. Heuristically, equation 3.13 shows that a weighted sum of the errors in each downstream layer times the derivative of the activation function for the neurons provides the generalized error for each hidden layer.

3.6 Considerations for Successful ANN's

The process of training and testing a successful ANN depends on many factors. Since the global error function may have many local minima, accurate training can be difficult. The manipulation of several constants used in the training process can help avoid problems during training.

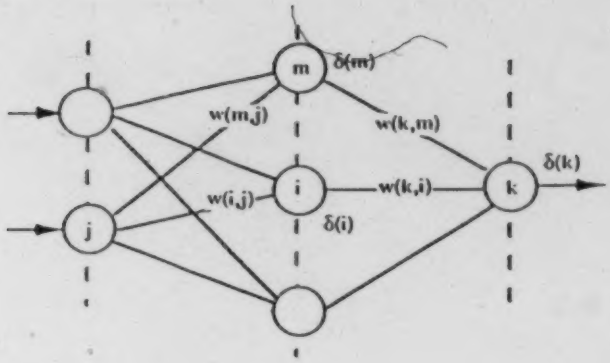


Figure 3.8: A three layer ANN with inter-neuron weights $w(i,j) \equiv w_{ij}$

By adjusting the learning parameter, η, the speed of the gradient descent may be adjusted. This allows the ANN to "jump" out of local minima by varying the weights w_{ij} enough to move the minimization on. A momentum value α may also be added to the learning process that provides a sort of memory as to how quickly the gradient descent was proceeding during the last iteration. If the last correction of the weights was large, the momentum term prevents the current step from becoming too small, and thus aids in the avoidance of local minima. We can write the effect of these two parameters on the evolution of the weights between interations n and $n-1$ as follows:

$$\Delta_n w_{ij} = -\eta \frac{\partial E_n}{\partial w_{ij}} + \alpha \Delta_{n-1} w_{ij}. \tag{3.14}$$

In addition to these parameters, a small amount of random noise can be applied to the weight evolution at each step in order to force the ANN out of local minima.

In practical applications, the training and evaluation of a feed-forward network

such as the ones used in this thesis consists of several steps. First, several variables that exhibit good discrimination for the classification task at hand are selected. The available set of known input events is then divided into three sets: the training set, the testing set, and the validation set. The input variables are scaled so that they all vary over the same range. This is important so that a given input node does not receive higher weights simply because its inputs are orders of magnitude higher than the others.

Training then proceeds by presenting a number of events from the training set to the network and allowing the weights to evolve through backpropagation. After a certain number of passes through the training set (each pass through the training set is called an *epoch*), the network output is evaluated on the events of the test set in order to gauge the success of the classification thus far. This procedure is repeated until either the performance of the network is deemed to be sufficient, or the change in the error function is unchanged over several epochs.

The *generalization performance* of the ANN is then evaluated using events from the validation set. These events have not been presented to the ANN before as they were not part of either the training or the test sets. If the validation set gives positive results, the ANN is ready for general use.

The basic parameters that affect the success of the ANN training are the size of the training event set, the number of nodes (neurons) in the ANN, and the number of epochs over which the ANN is trained. As a general rule, there should be at least several times as many events in the training set as there are weights in the network. This may be simply understood since the number of free parameters (represented by the ANN weights) must be much less than the number of available examples to fit those weights to. In general, the final generalization error of an ANN is on the order of the ratio of the number of weights to the number of training events. In order to combat this difficulty, the number of weights in a network can be reduced by removing some of the input variables in a process known as *pruning*.

The number of epochs over which to train is usually decided by watching the

performance of the network on the test set, and stopping when the error function on the test set fails to decrease any further. Going beyond the initial minimum of this error can be dangerous as the ANN may begin to model features in any unwanted noise in the data. When this happens, performance on the validation set degrades and the ANN is said to be *overtrained*. Overtraining can be guarded against by training the ANN for several different numbers of epochs and comparing the results on the validation set.

It is important to note that the classification performance given by an Artificial Neural Network will correspond to a Bayesian classifier if the training is accurate, only one output unit is switched on at a time, the mean square error function (equation 3.4) is used, and the ratios of the data of different types in the training sample matches the ratios in the generatization sets. In practice, achieving all of these conditions is difficult, but even in their absence, the ANN output still gives a useful classification variable.

3.7 Applications of ANN's in Particle Physics

The following sections will examine a few of the applications of Artificial Neural Networks in particle physics. Only the applications of feed forward networks for classification will be discussed, although it must be remembered that recurrent networks are increasingly finding application in such pattern recognition tasks as track finding.

3.7.1 The CDF Isolation Trigger

An excellent example of the power of an ANN in event classification may be found in the second level trigger of the CDF experiment at Fermilab [34]. This hardware implementation of an ANN is meant to tag events which may have an isolated electron in the electromagnetic calorimeter. Such an electron is indicative of W boson decay and hence is of much use in the physics analysis. Isolated leptons are also important

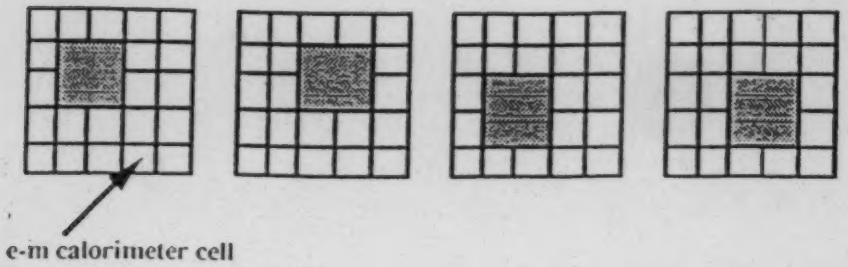


Figure 3.9: Selection of calorimeter clusters for the CDF isolation trigger ANN. The energy in each of the "inner" (shaded) areas is compared to the energy in the outer area of the 5x5 cell cluster in order to identify isolated leptons in the second level trigger.

to many Higgs Boson signatures.

For the purposes of this ANN, the cluster finding software identified the centre of an energy cluster in the 0.2 by 15° η/ϕ towers in the electromagnetic calorimeter. The energy content of each cell in the surrounding 5x5 block was then fed to the 25 input neural network (see figure 3.9). The network consisted of four hidden units each of which was trained to identify a structure similar to each of the templates in figure 3.9. The energy in the inner shaded area was compared to the energy in the remainder of the 5x5 cluster and if at least one of the four hidden neurons produced a value

$$f \cdot E_{inner} - E_{outer} > 0$$

then the isolation flag was considered to be true and an isolated electron was identified. From training of the network, the best value of f was found to be 0.16. Interestingly, this value of f agreed well with the most beneficial value found through trial and error offline analysis.

In practice, the implementation of this network in the second level trigger allowed a lowering in the energy threshold for electrons and reduced the background to W boson decays by a factor of four.

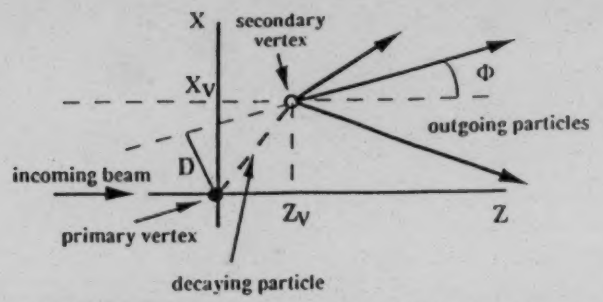


Figure 3.10: Geometric diagram showing how tracks originating from a common secondary vertex give rise to straight lines with the equation $D = R \cdot sin(\arctan \frac{x_v}{z_v} - \Phi)$ in D/Φ space.

3.7.2 Vertex Finding with Impact Parameters

Another interesting method for identification of vertices from heavy flavour particles uses the charged track parameters of impact parameter (D) and angle (ϕ) [35]. As can be seen in figure 3.10, tracks originating from the same secondary vertices form straight lines in D/ϕ space and the slope of these lines gives the decay distance R of the vertex from the primary vertex. Figure 3.11 shows an impact parameter versus angle plot for events in CDF. Note that secondary vertices are clearly visible as straight lines with non zero slope in the signal events while no such lines appear in the minimum bias background.

Of course, the limitations of detector resolution and noise severely curtail the ability of the analysis to pick out these straight lines in D/ϕ space. Attempts have been made, however, to use the impact parameter and initial track angles as input to a feed forward neural network in order to identify these straight lines. and therefore long lived vertices, in actual data. One such network consists of a feed forward ANN with 64 input units, 64 hidden units, and 8 output units. The input units were fed with the impact parameters and ϕ values of tracks in the central tracking mechanism. Figure 3.12 shows the identification efficiency versus cut on the neural network output for b vertices. One can see that the result is very good and in fact

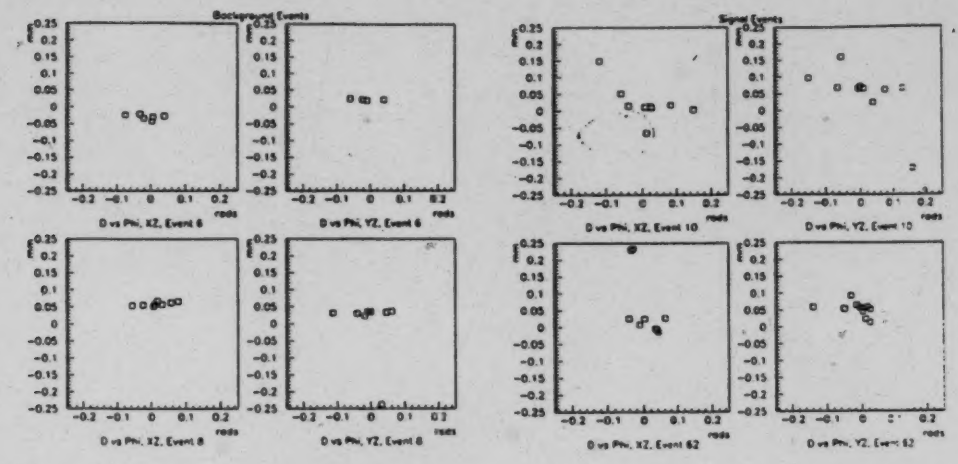


Figure 3.11: Impact parameter versus φ plots for events containing secondary vertices (right) and minimum bias background events (left) at CDF. Note the existence of identifiable straight lines in the left plot.

led to a factor of twenty reduction in minimum bias background contamination in the experiment in question.

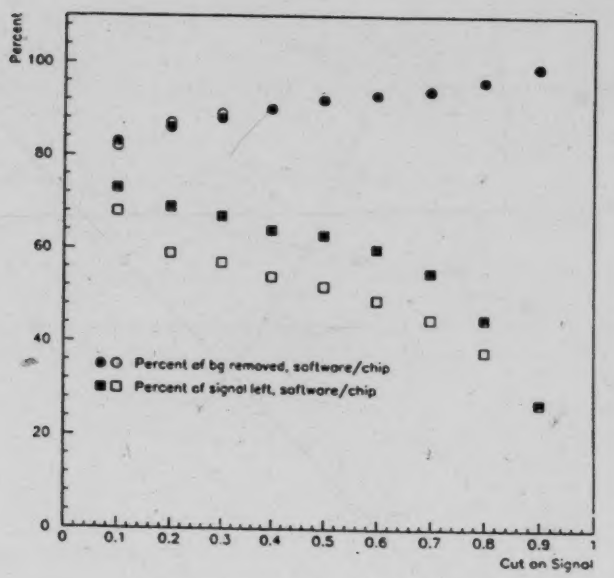


Figure 3.12: Identification efficiency of b vertices versus the cut made on the neural network output in CDF.

Chapter 4

The Search for a Neutral Higgs Boson

In this chapter, an Artificial Neural Network (ANN) based search for a neutral Higgs boson is presented. The analysis was developed to search for the Standard Model Higgs through the *four jet* channel $e^+e^- \rightarrow h^0_{SM}Z^0 \rightarrow \bar{b}b\bar{q}q$. As we saw in section 1.4.2, Standard Model Higgs analyses can be extended for use in the search for MSSM Higgs bosons produced through the Higgstrahlung process. Thus, the ANN analysis presented in this thesis has also been applied to the search for the lightest MSSM neutral scalar Higgs boson through $e^+e^- \rightarrow h^0Z^0 \rightarrow \bar{b}b\bar{q}q$.

In the sections that follow, the analysis will be developed in several stages. First, the overall strategy for the search will be defined. Then, the production of simulated data for ANN training will be described and the quality of the $\sqrt{s} = 183$ GeV data taken by OPAL in 1997 will be discussed. The offline analysis performed on the simulated and real data will then be presented and the discriminating variables for the analysis will be introduced. The ANN used for the analysis is then described and the training and testing procedure documented. Finally, the results of the analysis are applied to both the Standard Model and MSSM searches, and mass limits on the Higgs boson are derived.

4.1 Search Strategy

In order to define a successful search strategy for neutral Higgs bosons, we must look for differences in the signatures of the signal and background events. In this thesis, we search for neutral SM and MSSM Higgs bosons in the four-jet channel where the Higgs boson decays to $\bar{b}b$ quarks, and the Z^0 also decays to quarks. Since the strategy for MSSM Higgs detection is the same as the SM case, we shall hereafter refer only to the Standard Model Higgs search until we interpret the results in terms of the MSSM in section 4.8.

The signal events from $e^+e^- \rightarrow h^0_{SM}Z^0 \rightarrow \bar{b}b\bar{q}q$ exhibit four energetic jets with significant presence of heavy flavour due to the propensity for the Higgs to decay into b quark pairs. Since h^0_{SM} decays to b quarks roughly 80% of the time, and the Z^0 boson decays to quarks roughly 70% of the time, the four jet channel is by far the largest signature of the Higgstrahlung production mechanism at LEP. About 61% of all Higgstrahlung production results in the four jet final state at $\sqrt{s} = 183$ GeV.

The important background to the Higgs signal in both the SM and MSSM analyses consists of the following sources:

- Events from $(Z/\gamma)^* \rightarrow q\bar{q}(\gamma)$ where hard gluon emission mimics energetic jets.
- Four fermion decays from the production of W^+W^- and Z^0Z^0 pairs.
- Two photon events where gluon emission fakes quark jets.

With the exception of the irreduceable background $Z^0Z^0 \rightarrow b\bar{b}b\bar{b}$, all of these backgrounds can be reduced through tagging b quarks in the jets of the event.

Cross sections for these, and other, background processes are shown in figure 4.1 [7] with the production cross sections for the Standard Model Higgs for various Higgs masses. It may be seen in the figure that, for a centre of mass energy of 183 GeV, the cross section for the $(Z/\gamma)^*$ background is several orders of magnitude higher than the Higgs production mechanism while the W^+W^- and Z^0Z^0 cross sections are

of the same order as the signal. Thus, extracting a Higgs boson signal from the data is a challenging task that requires a large reduction in the background.

The general search strategy for the Higgs used in this thesis proceeds as follows. The particular aspects of each step will be explained in detail in the sections that follow.

1. Events are selected for the quality and for hadronic content in order to reject leptonic and poorly measured events.

2. The event is reconstructed into four jets and fit to the hypothesis of a Z^0 boson initiating one of the jet pairs. This fit is required to converge with good accuracy.

3. Cuts are made requiring well reconstructed and separated jets and a large amount of visible energy in order to reject Z/γ events with an initial state photon.

4. An Artificial Neural Network is applied that selects Higgs signal events based on further kinematic variables and the presence of b flavour in the jets of the event.

4.2 Production of Simulated Data

In order to train and test the Artificial Neural Network, and to evaluate the systematic errors present in the analysis, Monte Carlo simulations were performed to create data sets for both the signal and background channels. Signal datasets for $h^0_{SM}Z^0 \rightarrow b\bar{b}q\bar{q}$ were created using the generator HZHA [16]. This simulation used the parameters $\alpha_s = 0.118$, $m_t = 175$ GeV/c^2, $m_Z = 91.189$ GeV/c^2, and $G_F = 1.16639 \times 10^{-5}$ GeV^{-2}. The signal Monte Carlo sets used are listed in table 4.1. A large set of $m_h = 80$ GeV/c^2 data was generated for training and testing of the network while a set with Higgs mass ranging from 30 to 100 GeV/c^2 was generated for network validation.

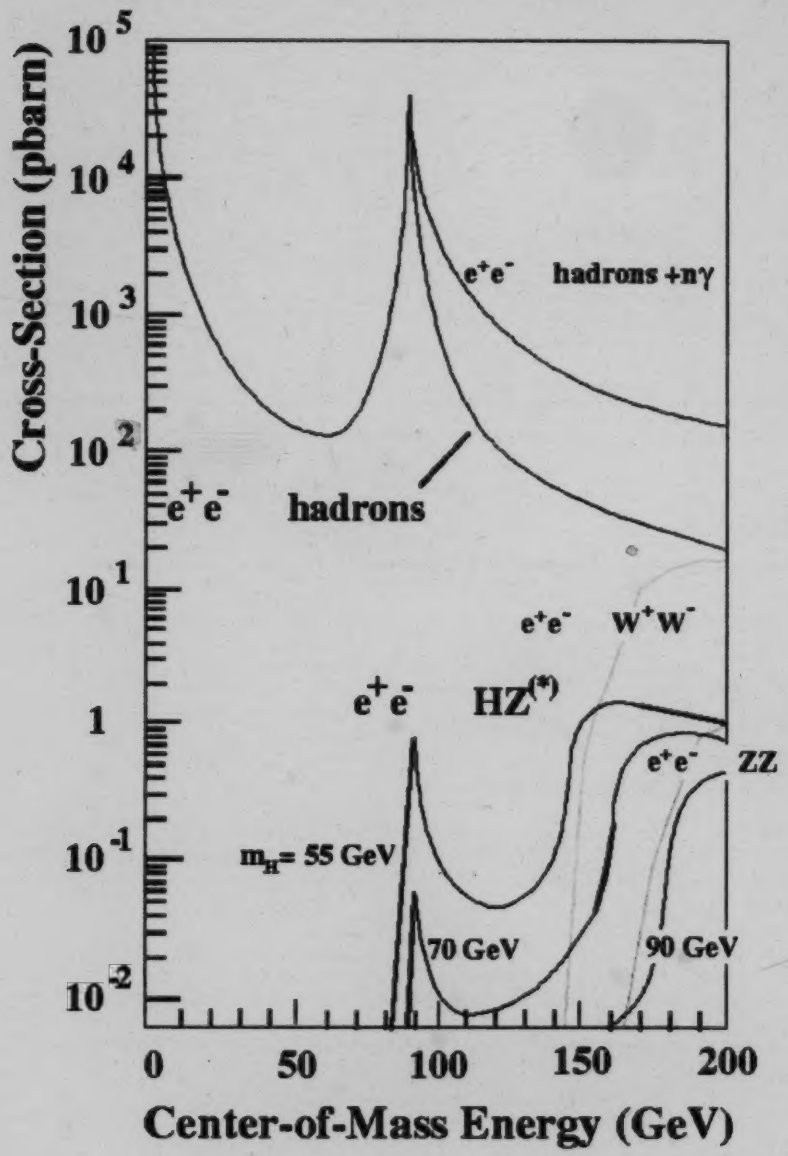


Figure 4.1: Cross sections for various background processes to the Higgs signal at LEP versus centre of mass energy as well as production cross sections for h^0_{SM} through the Higgstrahlung mechanism.

OPAL Run #	Process	Higgs Mass (GeV/c^2)	# Events	pb
6856	$e^+e^- \to h^0_{SM}Z^0$, $h^0_{SM} \to$ all, $Z^0 \to q\bar{q}$	80	5K	0.387
5769	$e^+e^- \to h^0_{SM}Z^0$, $h^0_{SM} \to$ all, $Z^0 \to q\bar{q}$	30 - 100	6K	varies
5050	$e^+e^- \to Z(\gamma)$	n/a	375K	107.43
7051	$e^+e^- \to q\bar{q}q\bar{q}$	n/a	30K	7.86
1026	$e^+e^- \to \gamma\gamma$	n/a	1000K	12.57 nb
1124	$e^+e^- \to Z(\gamma)$	n/a	150K	101.42
7330	$e^+e^- \to q\bar{q}q\bar{q}$	n/a	100K	18.171

Table 4.1: Monte Carlo data sets used in the search for the Higgs boson.

Background Monte Carlo (also listed in table 4.1) was also generated for the major background channels. For the (Z/γ) and two photon backgrounds, the PYTHIA [36] generator was used while four fermion final states including the W^+W^- and Z^0Z^0 backgrounds were generated with the GRC4F [37] generator. In order to later evaluate the systematic effects of using different Monte Carlo generators, the (Z/γ) background was also generated with the HERWIG [38] generator and the four fermion final states with the Excalibur [39] generator. All Monte Carlo events were passed through the OPAL detector simulation GOPAL [40] using the correct geometric description of the detector for the 1997 run.

4.3 The 183 GeV Data from OPAL

The data taken at $\sqrt{s} = 183$ GeV by OPAL in 1997 was collected in five "periods". The amount of data collected in each period is shown in table 4.2. During period 88, a set of three wires in the vertex chamber (CV) stereo layer broke. This malfunction, called the *CV incident* occurred at event 12773 in run 8679 [41]. No CV information was collected for an entire day, and only the axial portion alone was re-instated as of run 8692. By run 8816, two thirds of the CV stereo layers were working. Since the CV stereo information is relied upon to help match tracks from the jet chamber

Period	# Events	pb^{-1}
84	716	6.25
85	789	7.00
86	2593	22.84
88	1650	14.87
89	383	3.53

Table 4.2: The number events and integrated luminosities for OPAL data taken at $\sqrt{s} = 183$ GeV in 1997. Period 87 was taken at the Z^0 peak and so is not included in this analysis.

to the silicon microvertex detector in z, this mishap was potentially disastrous for b tagging. A software procedure was applied, however, that changed the algorithm for track matching between the jet chamber and silicon microvertex detector in such a way that the track matching efficiency was virtually unaffected after the incident. The track matching efficiency in z before and after the CV incident is shown in figure 4.2. The total useful integrated luminosity of data taken for the Higgs search in this thesis was measured to be 54.1 pb^{-1} by the ROCROS [42] analysis routine within OPAL.

4.4 Offline Data Analysis

In order to produce discriminating variables for the Higgs analysis, useful signatures must be calculated from the raw OPAL data. These signatures include jet properties, b tagging information, and the results of mass fits to the events. The sections that follow outline the variables calculated from OPAL data for use in this thesis.

4.4.1 Data Quality

In order to reject poorly measured events and events that obviously do not have the hadronic four jet nature desired in this analysis, a hadronic event selection was applied to all Monte Carlo and data events entering the analysis. In order to achieve this, cuts are made on reconstructed tracks and calorimeter clusters for each event.

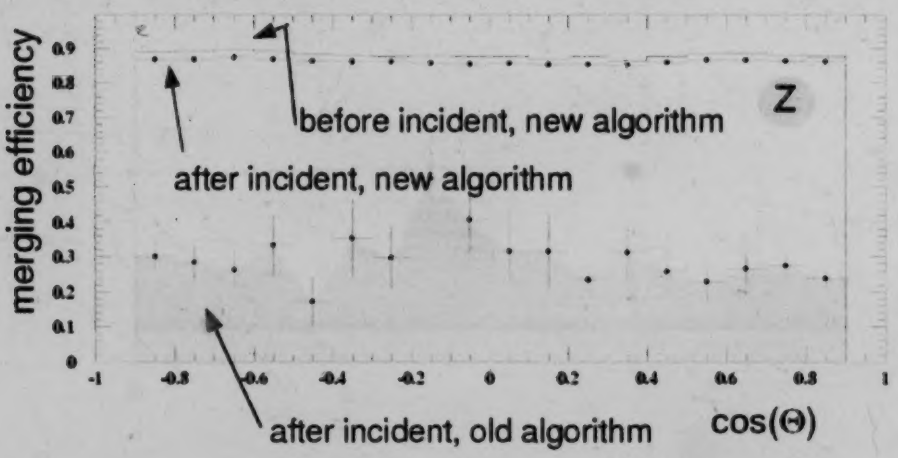


Figure 4.2: Merging efficiency for jet chamber and silicon microvertex tracks before and after the CV incident using the original and improved matching algorithms. With the improved matching algorithm, merging efficiency is virtually unaffected after the incident.

In order to understand these cuts, some track parameters in OPAL must be defined.

- d_0 (or b) is the track impact parameter which is defined as the distance from the origin to the point of closest approach of the track in the xy plane. The impact parameter is taken to be positive if the track, starting from the point of closest approach to the primary vertex, crosses the jet axis in the flight direction. Otherwise the impact parameter is negative.

- The dip angle of the track, λ, is given by $\tan\lambda = \cot\theta$.

- The value of the z coordinate at the point of closest approach of the track to the origin is given by z_0.

An initial cut on the data is made using the LEP2 Multihadron (L2MH) selection. This is a flag that is generated for each event in the offline analysis based on the

quality of tracks and clusters. In order for an event to be given an L2MH value of one, the following criteria must be fulfilled.

At least 5 tracks that pass the following OPAL [43] quality cuts:

- $0.050\,\text{GeV} \leq p_T$.
- The number of hits in the central detector > 20.
- $|d_0| \leq 2.0\,\text{cm}$.
- $|z_0| \leq 40.0\,\text{cm}$.
- Radius of the first hit in the central detector $\leq 60.0\,\text{cm}$.
- $|\cos\theta| \leq 0.995$.
- Chi squared of the $r\varphi$ fit to the track ≤ 999.
- Chi squared of the rz fit to the track ≤ 999.

and at least seven calorimeter clusters that pass the following cuts:

- The raw energy in each cluster is $\geq 0.100\,\text{GeV}$ in the barrel and $\geq 0.200\,\text{GeV}$ in the endcaps.
- The number of electromagnetic calorimeter blocks in the cluster is ≥ 1 in the barrel and ≥ 2 in the endcaps.

In addition to these track and cluster constraints, the event is required to have $R_{vis} = E_{vis}/2\sqrt{s} \geq 0.14$ where E_{vis} is the visible energy of the event, and $|R_{bal}| = |\sum E_{bal}/E_{raw}| \leq 0.75$ where $E_{bal} = E_{cluster}\cos\theta_{cluster}$ and E_{raw} is the raw energy in the event.

In real data, quality is also ensured through monitoring the OPAL detector status during data taking. For each OPAL subdetector, a *detector status bit* is recorded for each event that documents the state of the subdetector at that time. This status

Status Bit	Subdetector Status
0	Subdetector is dead
1	Subdetector is unreliable
2	Subdetector is experiencing small problems
3	Subdetector is reliable

Table 4.3: Detector status bits in OPAL.

Subdetector	Status Required
Silicon Microvertex	3
Central Vertex	2
Jet Chamber	3
Electromagnetic Barrel	3
Electromagnetic Endcap	3
Hadronic Calorimeter Towers	3

Table 4.4: Detector status bits required for real data in this analysis.

bit is an integer ranging from zero to three. Table 4.3 shows the definitions of the detector status bits in OPAL.

For the Higgs search presented in this thesis, it was important that the reliability of certain subdetectors be assured. To ensure accurate b quarks tagging, the central tracking detectors were required to be functional while good calorimeter status was also required to assure quality jet finding and energy measurements. Table 4.4 shows the status bits required for real data to be considered in this thesis. It is noteworthy that the central vertex (CV) chamber status bit was required to be 2 instead of 3 as runs during which only the axial part of the CV were operating after the CV incident were deemed to be of sufficient quality for analysis.

4.4.2 Jet Finding

The process of jet finding is of great importance in this analysis as the identification of four energetic jets is critical to the identification of signal events. To this end, the Durham [44] jet finder was used and was set to force each event to a four jet structure. If this procedure failed, the event was excluded from further consideration.

The Durham jet finder operates in a manner typical of cluster jet finders used in e^+e^- experiments. Initially, each identified particle is considered to be a cluster on its own. Then, the process of *binary joining* is begun with the distance between the two closest clusters being found with a certain metric. In the Durham jet finder, this metric is given by

$$\tilde{y}_{ij} = \frac{2min(E_i^2, E_j^2)(1 - \cos\theta_{ij})}{E_{cm}^2}$$

where E_i and E_j are the energies of the two clusters being considered, θ_{ij} is the angle between them, and E_{cm} is the centre of mass energy of the two clusters combined. If this distance is less than some threshold, the two individual clusters are joined into one. This procedure is repeated until all clusters in the event are separated by a distance larger than some predefined threshold and the remaining clusters are taken to be the jets in the event. This has the convenient result that each particle belongs to one and only one jet at the end of the binary joining.

The way the Durham jet finder is applied in this thesis stops the jet finding at just the point where four jets are defined and the distance measure at this point is recorded. Thus, we obtain a value of y_{34} which is the distance measure for which the event would change from a four jet to a three jet structure.

4.4.3 Energy Flow Correction

In OPAL, the energy of charged particles is counted twice: once in the central detector, and once in the calorimetry. In order to achieve the most accurate energy value for a given particle, and for the event as a whole, it is desirable to combine the energy information of the tracking and the calorimetry without double counting. In this analysis, the energy flow correction is performed using the OPAL Matching (MT) package [45]. This package aims to reduce the energy measured in a cluster with a matching track by an amount equal to the energy measured in the track. This is beneficial to the energy resolution of the experiment as the intrinsic energy resolution of the central tracking is superior to that of the calorimetry.

There are three possibilities for adjusting the track/cluster energy sharing within the MT processor. These are:

1. If a cluster has no associated track, both the energy of the cluster and the track information are used in the global energy of the event.

2. If a cluster is matched to a track and the energy in the cluster is less than expected given the energy measured in the central tracking, the cluster energy is ignored and only the track information is used.

3. If a cluster is matched to a track and the energy of the cluster exceeds the expected energy from the track information, the energy of the cluster is reduced by an amount equal to the expected energy deposition from the track. The track information and the reduced cluster energy are then both used in the event energy calculation.

4.4.4 Event Shape Variables

In order to quantify properties of the event as a whole, several *event shape* variables are derived from the OPAL data. These variables describe the spatial distribution of jets in the event. In order to extract these variables, we define the sphericity tensor:

$$S^{\alpha\beta} = \frac{\sum_i p_i^\alpha p_i^\beta}{\sum_i |p_i|^2} \tag{4.1}$$

where both α and β are indices for the three components of the momentum over all i particles in the event. This tensor is a real symmetric matrix, and hence can be diagonalized to obtain three eigenvalues $\lambda_i, i = 1, 2, 3$. Various combinations of these eigenvalues may then be used to define event shape variables.

The most common of these variables is the sphericity [46] which is defined by

$$S = \frac{3}{2}(\lambda_2 + \lambda_3).$$

The sphericity takes on values near one for isotropic (spherical) events and a value near zero for back to back two jet events.

In this thesis, the C parameter is used as the main event shape variable. The C parameter is also defined in terms of the eigenvalues of the sphericity tensor as

$$C_{par} = 3(\lambda_1\lambda_2 + \lambda_1\lambda_3 + \lambda_2\lambda_3). \tag{4.2}$$

The C parameter is a sensitive rejection tool for two jet events as it tends rapidly to zero for such event topologies. This is a very useful tool in background reduction for the Higgs search as the Z/γ background often has a two jet topology.

4.4.5 Effective Centre of Mass Energy

An important point regarding the Z/γ background at centre of mass energies above the Z^0 peak is that the production of initial state radiation (ISR) is favoured as it returns the hadronic energy of the event to the Z^0 pole through $e^+e^- \rightarrow Z^0\gamma \rightarrow q\bar{q}\gamma$. This *Z return* effect [7] allows the possible reduction of Z/γ background through the measurement of the *effective centre of mass energy* of the event. This value will be close to 91 GeV for Z return events, while it will be close to the centre of mass energy of 183 GeV for true four jet events.

In order to calculate the effective centre of mass energy, or $\sqrt{s'}$, the ISR photon must be identified and its contribution to the energy of the event subtracted. This is achieved by searching for an isolated photon cluster in the electromagnetic calorimeter with a cluster energy greater than 3 GeV. An isolated photon cluster must have the transverse shower shape expected of a photon as well as being isolated with no more than 1 GeV of energy in a cone of half-angle 200 mrad in the calorimeter. In about 20% of Z return events the ISR photon may be found in the calorimeter. In order to guard against photons that escape in the forward direction, a kinematic fit is also applied to the momentum and energy of the jets in the event with the added assumption that an unobserved photon escapes down the beam pipe. The final $\sqrt{s'}$ is calculated by reducing the LEP centre of mass energy by $2E_\gamma\sqrt{s}$.

4.4.6 B Jet Tagging

The identification of jets originating from B hadrons, and therefore from b quarks, is of the utmost importance to the analysis presented in this thesis as it allows the background from $W^+W^- \rightarrow q\bar{q}q\bar{q}$ and $Z^0Z^0 \rightarrow q\bar{q}q\bar{q}$ to be successfully rejected. Without b quark tagging, these backgrounds would be nearly impossible to control as kinematically they resemble the signal four jet events very strongly.

In OPAL, hadrons containing b quarks decay with decay lengths on the order of 2 mm. Thus, extrapolating the tracks resulting from a B hadron decay back to the interaction region with the silicon microvertex detector will reveal the existence of secondary vertices well separated from the primary e^+e^- interaction. As was discussed in section 2.2.2, the OPAL silicon microvertex detector is very well suited to finding such secondary vertices.

In addition to the presence of secondary vertices, jets originating from B hadrons can be identified through their higher charged track multiplicities and the presence of energetic leptons in the jet. Both of these signatures are due to the initial decay of the b quark via a virtual W particle which then increases the charged track multiplicities through hadronic decay, or creates the hard leptons through leptonic decay.

The process of b tagging at OPAL was developed centrally by the Higgs group [47] as its application is important to several channels in the Higgs search effort. The method is applied to each jet in an event and returns a b-tagging discriminator $\mathcal{B}_j$ for each of the j jets in the event. The scheme, shown schematically in figure 4.3, combines the results of a lifetime tag, a high-p_T lepton tag, and a jet shape tag in a Bayesian maximum likelihood estimator to produce the final jet tag.

The Lifetime Tag

The lifetime tag consists of five quantities that exploit the relatively long lifetime, high decay multiplicity, and high mass of B hadrons. The five quantities are combined in an artificial neural network (ANN) to determine the lifetime tag β_τ.

Three of the quantities depend on the reconstruction of secondary vertices in

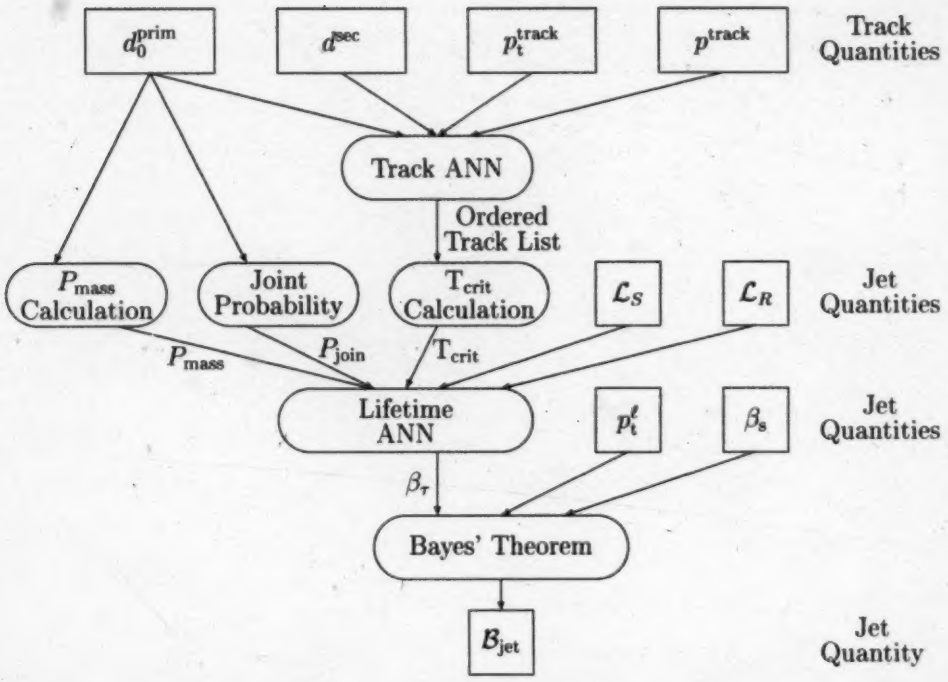


Figure 4.3: Schematic diagram of the procedure used to calculate the jet-wise b-tagging variable $\mathcal{B}_j$. The constituent parts are explained in the text.

OPAL. Within the jet being considered in the jet tag procedure, sub-jets are formed using a cone algorithm [48]. The cones in this algorithm have a half-angle of 0.5 radians and form sub-jets with a minimum energy of 7 GeV. Secondary vertices are identified in each of these sub-jets using the method described in [49] where it is attempted to fit all tracks in a given sub-jet to a single common secondary vertex. To do this, tracks are discarded in decreasing order of their contribution to the χ^2 of the secondary vertex fit until the largest χ^2 contribution is less than four. At least two tracks must remain to form a secondary vertex. This method of secondary vertex finding is sometimes called the "tear-down" method.

After performing this secondary vertex finding, the following three parameters are calculated for inclusion in the lifetime tag ANN.

1. **The Secondary Vertex Likelihood, $\mathcal{L}_S$**: This likelihood is calculated from the probability density function of S, the decay length significance[1], for b, c, and uds quark flavours. If more than one secondary vertex was formed in the sub-jet finding process, the $\mathcal{L}_S$ of the largest $\mathcal{L}_S$ secondary vertex is used in the ANN.

2. **The Reduced Secondary Vertex Likelihood, $\mathcal{L}_R$**: This likelihood is calculated in a manner analogous to $\mathcal{L}_S$, except the reduced decay significance is used. The reduced decay significance (R) is obtained by performing the secondary vertex fit without the track that has the largest impact parameter significance. For good b jet secondary vertices, R will not vary from S, but for random secondary vertices, the removal of the most significant track will often reduce the decay significance. If only two tracks make up a secondary vertex, R is not defined, and $\mathcal{L}_R$ is set to be equal to the relative likelihood for a b jet to form a two track secondary vertex compared to the likelihood for all jet flavours.

3. **Critical Track Discriminator, $\mathbf{T}_{\text{crit}}$**: To obtain the critical track discrim-

[1]The decay length significance is the decay length divided by the error in the decay length.

inator, a secondary artificial neural network is trained with four track-based inputs. These inputs are the impact parameter of the track with respect to the primary vertex, d_0^{prim}, the impact parameter of the track with respect to the secondary vertex, d_0^{sec}, the momentum of the track, p^{track}, and the transverse momentum of the track with respect to the sub-jet axis, p_t^{track}. The tracks in the sub-jet are then sorted in decreasing order of the secondary ANN output. A cluster of tracks is then formed by adding tracks one by one from the sorted list and calculating the invariant mass assuming pion masses for all tracks. The value of T_{crit} is defined as the output of the secondary ANN for the track which causes the invariant mass of the track cluster to exceed 1.9 GeV [50]. Since b flavoured hadrons have a higher mass than those formed of lighter quarks, this variable serves as a good discriminant.

Since not every b flavoured jet produces secondary vertices that can be detected in OPAL, two inputs to the lifetime tag ANN make use of track parameters only. These variables are

1. **The Two Dimensional Impact Parameter Joint Probability, P_{join}:** Tracks with negative impact parameters generally come from the primary vertex [50] and so the distribution of the impact parameters of such tracks gives an estimate of the resolution of the detector. This resolution is then used to weight the tracks and form the joint probability for tracks in the jet to come from the primary vertex [51]. This joint probability is given by

$$P_{\text{join}} = y \sum_{m=0}^{N-1} \frac{(\ln y)^m}{m!}$$

where y is the product of the probabilities of all N tracks in the jet to have come from the primary vertex.

2. **The Impact Parameter Mass Tag, P_{mass}:** The tracks in each sub-jet are sorted in descending order of impact parameter significance. In a manner

similar to the calculation of the Critical Track Discriminator above, the tracks are then iteratively clustered and P_{mass} is taken to be the impact parameter significance of the track that causes the invariant mass of the track cluster to exceed 1.2 GeV. The highest P_{mass} is used if a sub-jet contains more than one secondary vertex.

The five variables $\mathcal{L}_S$, $\mathcal{L}_R$, T_{crit}, P_{join}, and P_{mass} are then input into an artificial neural network based on the JETNET package [52]. The network has five input nodes (one for each input variable), a single hidden layer with ten nodes, and a single output node that provides the lifetime tag β_τ. The ANN was trained using Monte Carlo samples at $\sqrt{s} = m_{Z^0}$. Since the secondary vertex finding in OPAL is sensitive to the polar angle of the jets, three separate ANN's were employed with each covering a different polar region. These regions are $|\cos\theta| \leq 0.75$, $0.75 < |\cos\theta| < 0.9$, and $|\cos\theta| > 0.9$.

The High p_t Lepton Tag

In order to identify B hadron jets through the presence of energetic leptons, the standard OPAL electron and muon selections, which reject electrons from γ conversions [53], were applied. Once isolated energetic leptons are identified, their transverse momentum p_t^ℓ with respect to the axis of the sub-jet that includes the lepton track is used as an input to $\mathcal{B}_j$.

Jet Shape Tag

The jet shape is determined by calculating the sphericity (see section 4.4.4) of the jet in its own rest frame. This boosted sphericity, β_S, is closer to one for b jets due to their higher mass and decay multiplicity.

Combination of the Tags

Since the three tags discussed above exploit different aspects of the b jets, they are almost uncorrelated. In order to calculate a final b-tagging discriminant, $\mathcal{B}_j$, the

three tags are combined in an unbinned maximum likelihood. The final $\mathcal{B}_j$ for each jet is calculated as

$$\mathcal{B}_j = \frac{w_b f_b^\tau f_b^\ell f_b^s}{(w_b f_b^\tau f_b^\ell f_b^s) + (w_c f_c^\tau f_c^\ell f_c^s) + (w_{uds} f_{uds}^\tau f_{uds}^\ell f_{uds}^s)}$$

where w_b, w_c, and w_{uds} are weight factors that account for the different flavour composition of different background channels, f_q^τ are the probabilities for flavour q for the lifetime tag, f_q^ℓ are the probabilities for flavour q for the high p_t lepton tag, and f_q^s are the probabilities for flavour q for the jet shape tag. It has been found that the b-tagging sensitivity of $\mathcal{B}_j$ does not depend strongly on the choice of the weight factors. In this analysis, the weight factors are set to $w_b = w_c = w_{uds} = 1$.

4.5 Precut

Before selecting variables for an artificial neural network analysis, some loose precuts allow a large portion of the $(Z/\gamma) \rightarrow q\bar{q}$ background to be rejected. These precuts are designed to retain only true four jet hadronic events that have been well measured in the detector. The preselection cuts are as follows:

1. The L2MH flag as defined in section 4.4.1 must be on for the event.

2. The effective centre of mass energy $\sqrt{s'}$, as defined in section 4.4.5, must be greater than 120 GeV and the y_{34} parameter from the Durham jet finder must be greater than 0.003. The Durham jet finder is forced to find four jets.

3. Both a four parameter constrained fit which conserves energy and momentum and a five parameter constrained fit that adds a Z^0 mass hypothesis to one of the jet pairs must converge. This vetos poorly reconstructed events.

4. Each of the four reconstructed jets must contain at least two charged tracks.

Figure 4.4 shows the distribution of $\sqrt{s'}$ and y_{34} after the L2MH cut for Monte Carlo 80 GeV/c^2 Higgs events and for (Z/γ) background. The strong rejection power

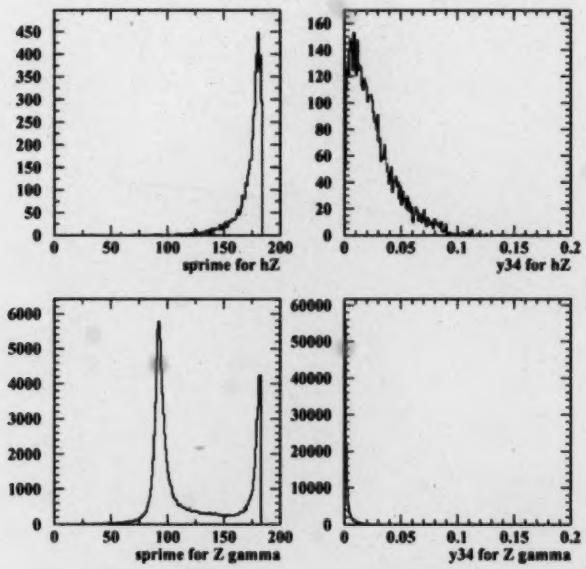


Figure 4.4: Distributions of $\sqrt{s'}$ and y_{34} for simulated 80 GeV/c^2 Higgs events (top row) and (Z/γ) background (bottom row). The rejection power of the precuts is evident.

of the $\sqrt{s'}$ and y_{34} cuts is evident. Figure 4.5 shows the same comparison between 80 GeV/c^2 Higgs events and four fermion final state events. The preselection cuts in this case have much less effect, and other measures must be taken to reduce the backgrounds.

Table 4.5 shows the number of remaining 183 GeV data events after each of the four precut steps as well as the amount of remaining Monte Carlo for each of the background channels. The number of background Monte Carlo events have been normalized to the number of data events for easy comparison. We see that the precuts are very effective at reducing the $(Z/\gamma) \rightarrow q\bar{q}\gamma$ background while keeping the efficiency for the Higgs signal high. The agreement between the remaining number of data events and the remaining number of Monte Carlo events is also good at each stage of the cuts.

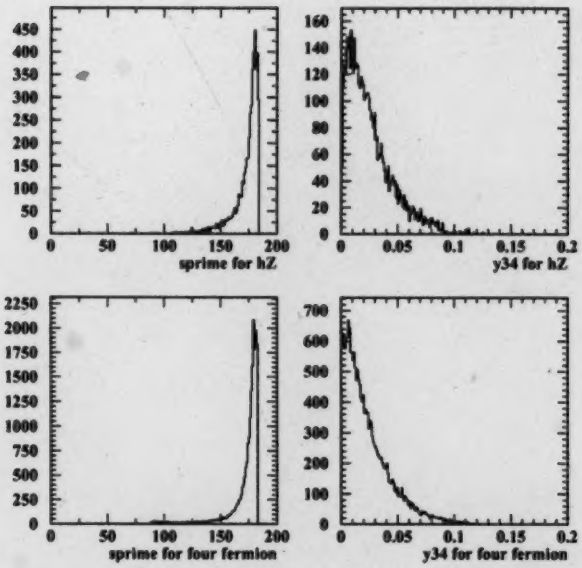


Figure 4.5: Distributions of $\sqrt{s'}$ and y_{34} for simulated 80 GeV/c^2 Higgs events (top row) and four fermion final state background (bottom row). The precuts do not have much rejection power in this case.

Cut	Data 183 GeV	Total Bkg.	$q\bar{q}(\gamma)$	4f	2 photon	Efficiency (%) $m_{h^0} = 85$ GeV
1	6131	6118.5	5099.0	948.4	71.1	99.8
2	941	907.9	401.6	499.8	6.5	94.2
3	779	747.3	317.9	422.9	6.5	93.0
4	632	607.4	229.6	372.4	5.4	86.1

Table 4.5: The number of events surviving each precut for the 183 GeV data and the expected background normalized to 54.1 pb^{-1}. Efficiencies for the signal $Z^0 h^0_{SM} \rightarrow q\bar{q}b\bar{b}$ with a Higgs mass of 85 GeV are shown in the last column.

4.6 The Higgs Search ANN

In this thesis, an artificial neural network (ANN) technique is used to classify events as either $h^0_{SM}Z^0 \to q\bar{q}b\bar{b}$ signal, $Z^0/\gamma^* \to q\bar{q}$ background, or four-fermion background after the preselection. The ANN is of the standard back-propogation type presented in chapter 3 and is based on the JETNET package [52] version 3.5. The network consists of three layers with ten input nodes in the first layer, 11 nodes in a single hidden layer, and three nodes in the output layer. All of the nodes are dot product neurons with sigmoidal activation functions. Each of the input nodes corresponds to one of the input variables which are described below.

4.6.1 Input Variables

The ten input variables selected for this analysis were chosen for their potential to discriminate between the signal and background signatures. Variables were chosen to discriminate against either the $Z^0/\gamma^* \to q\bar{q}$ or the four fermion background. Seven of the input variables exploit the kinematics of the event while three make direct use of the b-tagging discriminants $\mathcal{B}_j$ described in section 4.4.6. The ten input variables are as follows:

1. **Emax-Emin**: The difference in energy between the highest and lowest energy jets in the event. This variable helps separate signal and $Z^0/\gamma^* \to q\bar{q}$ background since gluon jets are generally less "hard" than quark jets.

2. **m(dijet)min**: The invariant mass of the lowest mass jet pair in the event. This variable also discriminates against $Z^0/\gamma^* \to q\bar{q}$ background as jet pairs including gluon jets will have a lower invariant mass than jet pairs including hard quark jets.

3. **sprime**: The effective centre of mass energy, $\sqrt{s'}$, of the event. As defined in section 4.4.5, this variable is effective in reducing the $Z^0/\gamma^* \to q\bar{q}$ background.

4. **C-parameter**: The C parameter of the event as defined in section 4.4.4. This variable reduces the background due to two jet like events from $Z^0/\gamma^* \rightarrow q\bar{q}$.

5. **Betamax**: The dijet pairing closest to the W pair hypothesis is selected that minimizes $(p_{\text{dijet1}} - 45\ GeV)^2 + (p_{\text{dijet2}} - 45\ GeV)^2$. The input variable β_{max} is then taken to be the larger of $\beta = p_{\text{dijet}}/E_{\text{dijet}}$ for the selected pairing. This variable helps to reduce the four fermion background from W pairs.

6. **Prob6c**: The probability of a constrained fit to the event that enforces momentum and energy conservation (four constraints) as well as the hypothesis that each of the two dijets in the event come from a W (2 constraints). This variable helps to reject W pair events.

7. **Prob5c**: The probability of a constrained fit to the event that enforces momentum and energy conservation as well as the hypothesis that one of the two dijets in the event comes from a Z^0. This is the signal hypothesis, and this variable helps to reject events where a good Z^0 candidate cannot be reconstructed.

8. **Bdisc1**: The b discriminant $\mathcal{B}_j$ for the jet with the highest $\mathcal{B}_j$ in the event. This is a general b tagging variable that helps to reduce all backgrounds to the Higgs signal.

9. **Bdisc2**: The b discriminant $\mathcal{B}_j$ for the jet with the second highest $\mathcal{B}_j$ in the event.

10. **Bdisc3**: The b discriminant $\mathcal{B}_j$ for the jet with the third highest $\mathcal{B}_j$ in the event.

Figure 4.6 shows the distributions of the ten input variables for OPAL data, total background Monte Carlo, and simulated Higgs events. It may be seen from the figure that the agreement between background Monte Carlo and OPAL data is very good for all of the input variables. The discriminating power of the variables may also be seen from the difference in shape between the signal and background distributions.

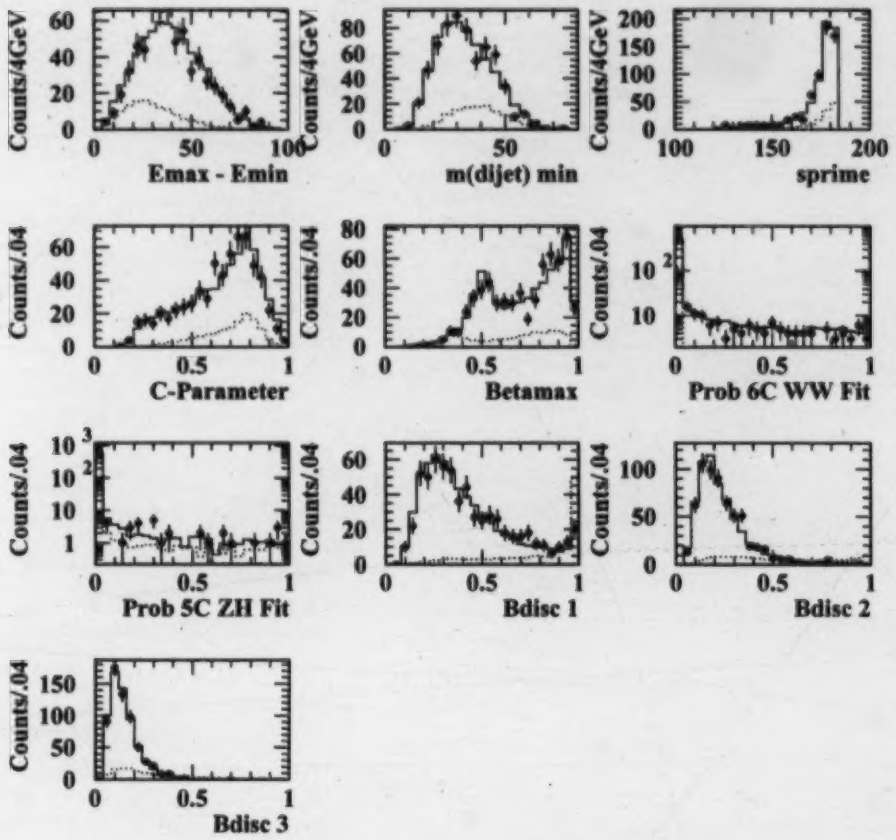


Figure 4.6: The distributions of the ten input variables to the Higgs search ANN for OPAL data (points), the sum of background Monte Carlo (solid histogram), and signal × 10 for a simulated $m_{h^0_{SM}} = 85$ GeV/c^2 sample. All distributions are taken after the precuts have been applied.

4.6.2 ANN Training

The ANN was trained with 4000 $h^0_{SM}Z^0 \to b\bar{b}q\bar{q}$ events, 8000 (Z/γ) events, and 4000 four fermion final state events. During training, the three output nodes were given a target of (1,0,0) for signal events, (0,1,0) for (Z/γ) background, and (0,0,1) for four fermion background. After every ten training epochs, the network was tested on a different set of 4000 signal events, 8000 (Z/γ) events, and 4000 four fermion events. All events used in the training and testing of the ANN had passed the precuts found in section 4.5.

The following parameters were set to govern the learning of the ANN:

- The learning parameter, η, was set at 0.001.
- The momentum, α, was set at 0.5.
- The temperature, which controls the slope of the sigmoid activation function, was set at 1.
- The inter-node weights were updated for each pattern processed.

As discussed in chapter 3, it is important to present an ANN with input variables that all range over the same order of magnitude. To this end, the input variables were scaled to a mean of zero and unit variance before being presented to the ANN.

The network was trained for 500 epochs and then stopped. Figure 4.7 shows the mean squared error versus the training epoch for both the training and the testing sets. It may be seen from the figure that the error finds a stable minimum before 500 epochs and that there is no sign of overtraining in the test set error. If overtraining were to occur, the test set error would begin to rise while the training set error continued to decrease. The network was trained out to 10,000 epochs for the purposes of testing overtraining, but none was observed.

Figure 4.7 also shows the mean saturation for the nodes of the hidden layer, and the mean saturation for the output layer. The saturation is a measure of how

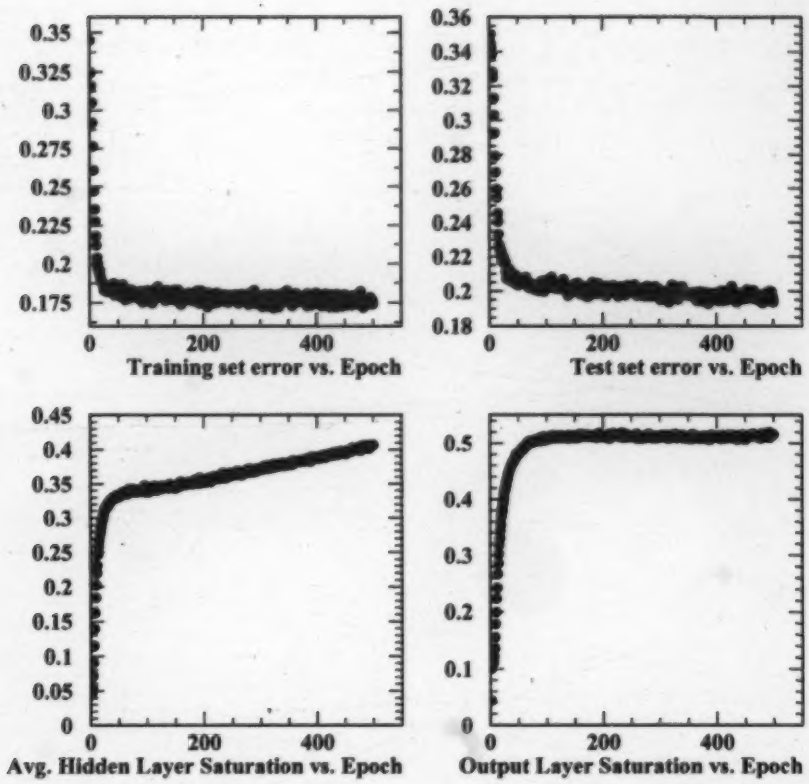


Figure 4.7: Mean squared error for the training (top left) and test (top right) sets versus epoch for the ANN training. The bottom two plots show the mean saturation of the nodes of the hidden layer (left) and the mean saturation of the nodes of the output layer (right). The developing plateau in both of these plots indicates that the network has learned its full capacity of information.

far the neurons have evolved towards the extreme ends of the activation function. Thus, a plateau in the saturation is indicative of a network that has "learned" its full capacity. We see from the figure that this is indeed the case with our ANN.

The efficiency evolution of the ANN for the Higgs signal is shown versus training epoch in figure 4.8. We can see that the training is successful and that the efficiency levels out well after climbing rapidly early in the training process. Figure 4.9 shows the distributions of the three output nodes of the network for simulated 85 GeV/c^2 Higgs signal and for the two major backgrounds: (Z/γ) events and four fermion events. It may be seen in the figure that the network successfully classifies the three event types. The first output node peaks at one for the signal events and is close to zero for both background types, the second node peaks near one for (Z/γ) events and is close to zero for the other classes, and the third node peaks near one for the four fermion events while being close to zero for other event classes.

4.7 Results of the Standard Model Search

In order to match the background expectation of the ANN analysis developed in this thesis with other analysis methods used in OPAL for comparison purposes (see section 4.7.2), a final ANN cut at a value of 0.85 for the first output node was decided upon. Thus, the final analysis consisted of the precuts presented in section 4.5 followed by the ANN first node cut at 0.85. The results of the analysis on the OPAL 183 GeV data, the background Monte Carlo, and the signal Monte Carlo are shown in table 4.6. In the 54 pb^{-1} of data collected by OPAL at $\sqrt{s}$ = 183 GeV, 6.15$\pm$0.15 (stat.) events due to background are expected while 7 events are observed. The Monte Carlo indicates that 2.7 of these background events are (Z/γ) events while 3.4 are four fermion final states. The two photon background is completely rejected by the ANN. The efficiency of the analysis on simulated Higgs events of different masses is shown in table 4.7. The efficiency peaks at a value of 37.2% for Higgs masses of 85 and 90 GeV/c^2. The efficiency curve is relatively flat in the region

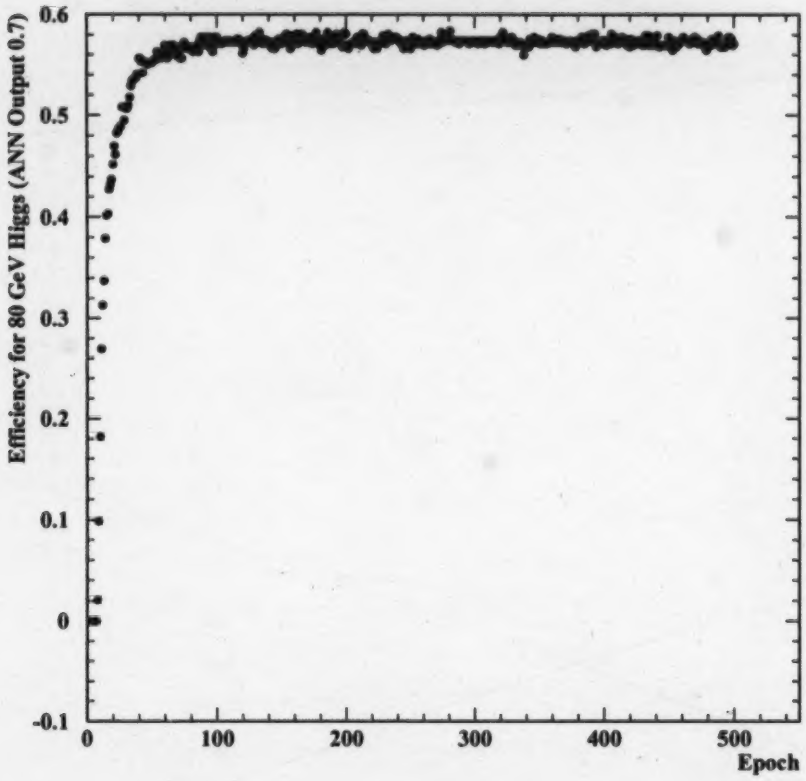


Figure 4.8: Efficiency versus training epoch for 80 GeV/c^2 Higgs events in the test set. The event was selected as signal if the ANN output was greater than 0.7.

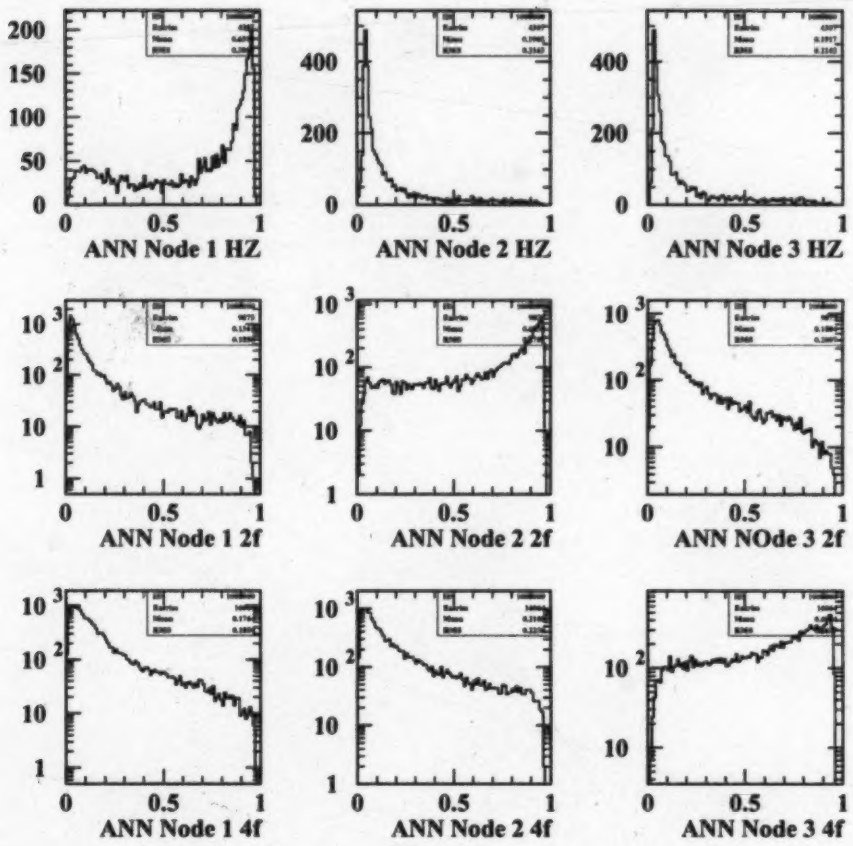


Figure 4.9: Output of the three ANN output nodes (columns) for the signal (first row), Z/γ^* background (second row), and four fermion background (third row).

Cut	Data 183 GeV	Total Bkg.	$q\bar{q}(\gamma)$	4f	2 photon	Efficiency (%) $m_{h^0} = 85$ GeV
Precut	632	607.4	229.6	372.4	5.4	86.1
ANN > 0.85	7	6.15 ± 0.52	2.7	3.4	0.0	37.2

Table 4.6: The number of events surviving the precut and the subsequent ANN cut for the 183 GeV data and the expected background normalized to 54.1 pb^{-1}. Efficiencies for the signal $Z^0H^0 \rightarrow q\bar{q}b\bar{b}$ with a Higgs mass of 85 GeV are shown in the last column.

m_H (GeV)	Efficiency $h^0_{SM} \rightarrow b\bar{b}$(%)
30	12.7
40	16.0
50	23.1
60	26.8
65	25.0
70	32.0
75	33.8
80	35.4
85	37.2
90	37.2
95	35.5
100	29.9

Table 4.7: Efficiencies for $H^0Z^0 \rightarrow b\bar{b}q\bar{q}$ vs. Higgs mass for the ANN four jet analysis.

$70\,\text{GeV} \leq m_{h^0_{SM}} \leq 95\,\text{GeV}$. For an 85 GeV/$c^2$ Higgs boson, the efficiency of the analysis corresponds to an expected signal of 6.44 events.

Figure 4.10 shows the output of the first ANN node for each of the background channels, the total expected background, and total background plus the expected signal from 85 GeV/c^2 Higgs events. The ANN Output of OPAL data is also presented in the figure and it can be seen to agree well with the total background hypothesis. The seven OPAL data events that pass the ANN cut can be seen in the figure. The first node ANN outputs and the reconstructed Higgs masses[2] may be found in table 4.8. An event picture of one of these candidates is shown in figure 4.11.

[2]The Higgs mass is reconstructed using the invariant mass of the jet pair opposite the dijet assigned to the Z^0 in the 5 parameter Z^0 mass constrained fit.

Run	Event	ANN First Node Output	$m_{h^0_{SM}}$ (GeV)
8138	172786	0.866	52.5
8221	34856	0.924	88.8
8284	28362	0.900	82.9
8422	57130	0.920	85.0
8712	89200	0.902	75.6
8812	21472	0.950	67.9
8820	60326	0.883	89.1

Table 4.8: Candidate events for the ANN four jet analysis. The ANN output and the reconstructed Higgs mass for each candidate is shown.

The four jet structure of the event is clear.

4.7.1 Evaluation of Systematic Errors

Several systematic checks were performed on the ANN analysis in order to test its sensitivity to changes in the input variables or detector performance. The uncertainty in the track resolution due to imperfections in the Monte Carlo was estimated by smearing the impact parameter resolution in the Monte Carlo by $\pm 5\%$ in $r\varphi$ and $\pm 10\%$ in z. Comparing the number of accepted signal and background events after the smearing to the nominal ANN analysis gave a relative percent error of 1.6% in the signal[3] and 6.1% in the background.

Uncertainties in the simulation of B hadron fragmentation were estimated by varying the B hadron charged multiplicity by ± 0.35 and the mean value of the B hadron energy spectrum by ± 400 MeV [54]. Changing the charged multiplicity resulted in a relative error of 0.8% in the signal and 3.6% in the background. Shifting the B hadron energy spectrum gave a 4.8% relative error in the signal and 4.6% in the background, thus showing the sensitivity of the B tagging variables used in the ANN to this systematic.

The modeling of the background events was tested by evaluating the performance of the ANN analysis on background data sets created with different Monte Carlo

[3]The 85 GeV/c^2 Higgs sample was used to evaluate the systematic errors in the signal.

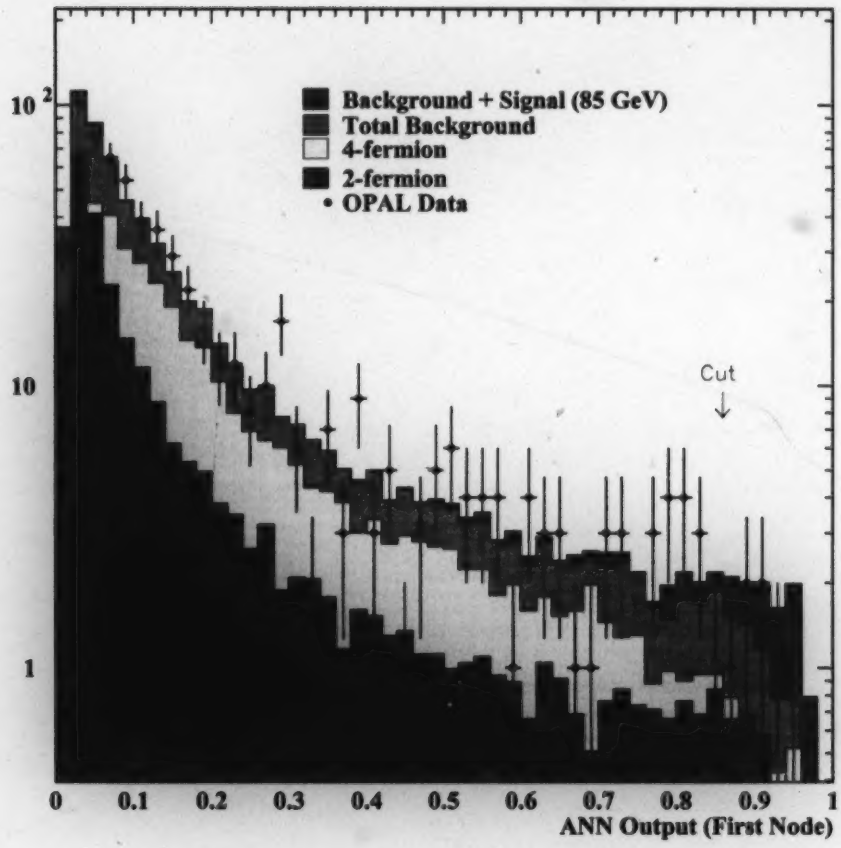


Figure 4.10: Output of the first ANN node for the background channels, the total background, and the total background plus an amount of simulated 85 GeV Higgs signal normalized to the integrated luminosity. The ANN output of OPAL data is also shown by the points. The cut on the ANN output at 0.85 is shown and the enhancement of the signal above the cut is visible.

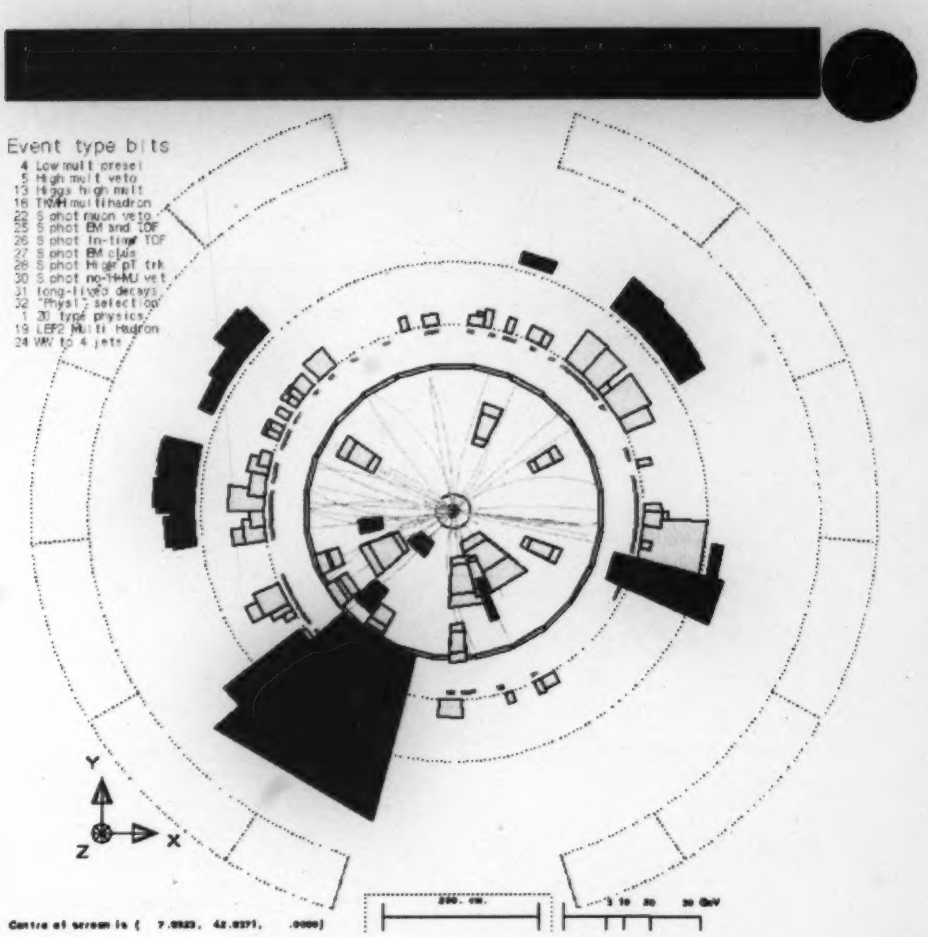


Figure 4.11: A candidate event selected by the Artificial Neural Network selection used in this thesis. The ANN output for this event was 0.9 and the reconstructed Higgs mass was 82.9 GeV/c^2.

generators. The ANN was applied to (Z/γ) events generated with the HERWIG [38] generator and four fermion events generated with the EXCALIBUR [39] generator. The difference between the number of events expected with the alternative Monte Carlo generators and the baseline analysis contributed a 2.9% relative error to the background estimation.

A transformation method was used in order to account for the modelling accuracy of the input variable distributions in the Monte Carlo. This method makes use of a fractional probability integral in order to create a transformation between the original (Monte Carlo) distribution and the target (data) distribution [55]. The original and target distributions are first smoothed into continuous functions using a spline fit and then the fractional integral is used to create a mapping between the distributions. This mapping is then applied to the Monte Carlo input distributions before evaluating them using the ANN, and the resulting change in the number of accepted events is quoted as a systematic error. Applying the transformation mapping to the signal Monte Carlo and then applying the ANN results in a relative uncertainty of 5.9%. Performing the same analysis with the background Monte Carlo results in a relative error of 17.7%.

Since the transformation method takes into account errors implicit in the impact parameter smearing, only the former is taken into account in the total systematic error for the signal and background. The total systematic error is calculated by adding the above sources in quadrature and results in a relative percent error of 7.6% for the signal and 18.9% for the background. Statistical errors due to Monte Carlo statistics are estimated to be 0.7% for the signal and 8.5% for the background.

To summarize, then, the expected background for the Standard Model Higgs analysis in the four jet channel presented in this thesis is 6.15 ± 0.52(stat.)± 1.16(syst.) events. Seven events are observed in data thus matching well with the expected background. If a Higgs boson were to exist with a mass of 85 GeV/c^2, the ANN analysis would detect it with an efficiency of 37.2 ± 1.0(stat.) $\pm$ 2.9(syst.)%

4.7.2 Comparison with Other Analysis Methods

A maximum likelihood approach was also employed by the Bonn OPAL group for the Standard Model Higgs search in the four jet channel [47]. This analysis uses stronger precuts followed by an unbinned maximum likelihood analysis using similar variables to those used in the ANN analysis developed in this thesis. The efficiencies of the two analyses are very similar with the likelihood analysis having a peak efficiency of 39.2% for a 85 GeV/c^2 Higgs. The systematic errors of the two analyses are comparable in magnitude.

The same Monte Carlo signal and background sets were used by the ANN analysis and the maximum likelihood analysis. Over 80% of the selected signal Monte Carlo events were in common between the two analyses while 60% of the selected background Monte Carlo was shared between the two. This is consistent with the fact that five of the seven selected candidates in the 183 GeV OPAL data were common to the two analyses.

4.7.3 Mass Limits for the SM Higgs Boson

Mass limits for the Standard Model Higgs boson are calculated centrally by the Higgs group at CERN in order to combine information from all relevant search channels. Each of the LEP experiments uses its own statistical method to calculate the confidence levels for mass limits. The OPAL procedure is described in the section that follows.

A general confidence limit scheme begins by defining a test-statistic[4], X, from the number and pattern of events selected by the analysis. The observed test-statistic, X_{obs}, is compared to the distributions of the same test-statistic generated on the basis of "thought experiments" for particular Higgs boson mass hypotheses. A given mass hypothesis is excluded at the confidence level $1 - CL$ if CL is the probability

[4] A *statistic* is a function of observables (such as the reconstructed Higgs mass) of a population sample and is independent of population parameters. In contrast, a *test-statistic* is a function of both the observables and the parameters (e.g. the *true* Higgs boson mass) which is used in hypothesis testing.

for "thought experiments" to obtain $X_{s+b} < X_{obs}$ where X_{s+b} is the test-statistic expected for the *signal and background* hypothesis. This is the classical frequentist approach described by

$$CL_{s+b} = P(X_{s+b} < X_{obs}). \tag{4.3}$$

The classical frequentist approach is not perfect since it does not consider the possibility that a particular observation might be such that even the background, with no signal present, is excluded. To rectify this problem, OPAL uses, instead of CL_{s+b}, the normalized signal confidence level [4]

$$CL_s = \frac{P(X_{s+b} < X_{obs})}{P(X_b < X_{obs})} \tag{4.4}$$

where X_b is the test-statistic expected assuming background only and X_s is the test-statistic assuming signal only. The 95% confidence limit for the Higgs boson mass is obtained for the hypothetical mass value that yields $CL_s = 0.05$.

The exact statistical method used in OPAL to quantify equation 4.4 is based on fraction event counting [56]. The test-statistic is a sum of weights over candidate events in all channels: $w = \sum_i w_i$. The w_i are calculated according to a *filter function* $w_i = F(m_i)$ where m_i are the reconstructed Higgs boson masses in the event.

The filter function, $F(m_i)$, depends on the hypothetical Higgs boson mass, m_H. Therefore, the total weight w is also a function of m_H and becomes large at values of m_H where candidate events cluster. The probability distribution of the total weight, $P(w)_{s+b}$, including signal and background events, forms the basis for the evaluation of confidence levels.

In principle, the filter function is arbitrary. The construction of the weight distribution function, $P_1(w)$, for the case of one signal event and no background event, is illustrated in figure 4.12, assuming a specific form for $F(m)$. In the figure, the function $D(m)$ describes the actual mass distribution and the hatched areas indicate the probability $P_1(w)\Delta w$ for a Higgs boson to yield a weight inside the interval Δw. The $P(w)$ distributions for fixed numbers of signal events and no background event

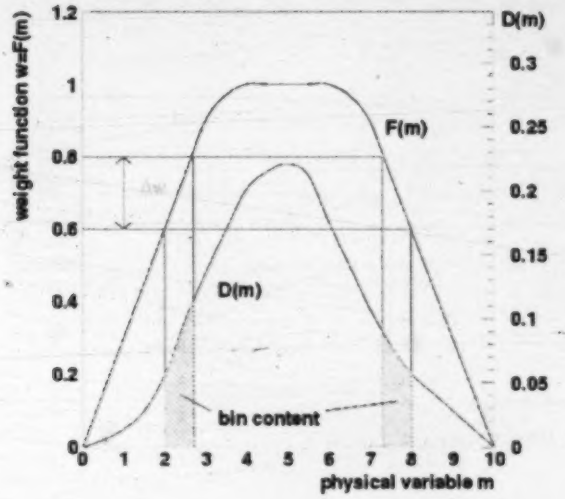


Figure 4.12: Relation between the mass distribution $D(m)$, the filter function $F(m)$, and the probability $P_1(w)\Delta w$ (hatched area) for one candidate event to have a weight $w \pm \Delta w/2$. In the example chosen $w \pm \Delta w/2 = 0.7 \pm 0.1$.

can be obtained from $P_1(w)$ by iterative integration, and the total probability density $P_s(w)$ is obtained by adding the whole set of functions, weighted with the Poisson distribution for the expected number of Higgs bosons. The distribution $P_s(w)$ is thus sensitive to the Higgs boson production rate.

An analogous procedure leads to the distribution $P_b(w)$ of background events which can be combined with the signal distribution $P_s(w)$ to evaluate the overall distribution $P_{s+b}(w)$. The function $P_{s+b}(w)$ then includes the background predicted by the Standard Model and depends on the signal rate which is a function of m_H.

The confidence level may then be computed, in accordance with equation 4.4 as

$$CL_s = \frac{\int_0^{w_{obs}} P_{s+b}(w)dw}{\int_0^{w_{obs}} p_b(w)dw}. \tag{4.5}$$

The denominator is introduced to avoid situations with excessively low limits arising when less candidates are observed than expected.

The filter function used in OPAL is chosen to provide good discrimination between signal and background events. For a single decay channel and a flat background distribution, it is adequate to use the mass distribution itself: $F(m) = \frac{D(m)}{D_{max}}$ where D_{max} is the maximum value of $D(m)$. When several channels with different efficiencies and backgrounds are to be combined, however, a channel-dependent filter function is introduced to account for the different signal to background ratios:

$$F_k(m) = K \cdot \frac{1}{C + \frac{b_k(m) \cdot \sum_k S_k}{D_{k,max} \cdot S_k}} \cdot \frac{D_k(m)}{d_{k,max}}. \tag{4.6}$$

where S_k is the expected signal rate in channel k and b_k is the differential background rate. The factor K is chosen so as to fix the largest value of $F_k(m)$ to unity. The constant C is a free parameter which is optimized by minimizing the signal confidence level in background-only thought experiments as a function of the Higgs boson mass hypothesis.

The OPAL filter function in equation 4.6 has the following properties:

- In the limit of very low background, $F_k(m)$ approaches a universal value K/C at the mass peak positions independently of the channel number k.
- $F_k(m)$ is invariant under arbitrary splitting of any decay channel.
- If the background is large compared to the signal, $F_k(m)$ is proportional to the local signal to background ratio.
- For C equal to the total number of expected events, the maximum of $F_k(m)$ is proportional to the probability for a candidate in channel k to be a signal event.

Merging the ANN analysis developed in this thesis with the other Standard Model Higgs search channels in OPAL gives a 95% confidence level lower mass limit of 86.94 GeV/c^2 for h^0_{SM} following the above prescription. This is a significant improvement over the pre-183 GeV mass limit of 69.4 GeV/c^2.

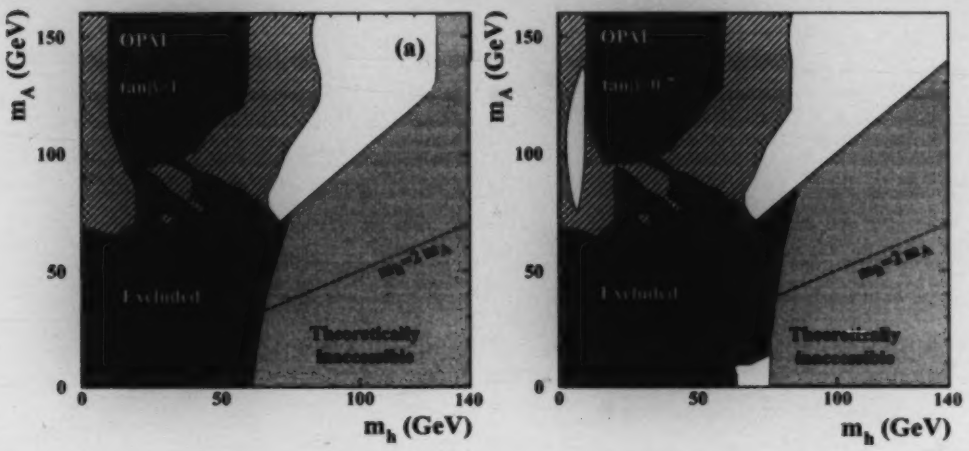


Figure 4.13: MSSM exclusion plots for $\sqrt{s} \leq 183$ GeV. Excluded regions are shown in the (m_{h^0} , m_{A^0}) plane at the 95% confidence level. The left plot shows the result for $\tan\beta > 1$ and the right plot for $\tan\beta > 0.7$. The black areas are excluded solely by comparing data to the expected MSSM production rates, without applying any additional theoretical criteria. The speckled grey areas are theoretically inaccessible.

4.8 Results of the MSSM Search

As we saw in section 1.4.2, for values of $\tan\beta > 0.7$, we can apply Standard Model analyses directly to the search for the lightest neutral scalar MSSM Higgs boson, h^0. The OPAL MSSM general scan collects the efficiency and background spectra for each relevant analysis and combines them in a confidence level analysis (see section 4.7.3) to produce exclusion plots. When the four jet analysis developed in this thesis is added to the analyses used in the OPAL MSSM general scan, the exclusion plot shown in figure 4.13 is the result. For $\tan\beta > 1$, the combined analysis excludes the CP-odd neutral MSSM Higgs boson at masses below 64.5 GeV. For $\tan\beta > 0.7$, the region $m_{A^0} < 13\,\text{GeV}$ is not excluded. Applying additional theoretical constraints to the model [47] allows the 95% confidence limit exclusions to be extended to $m_{A^0} > 67.5\,\text{GeV}$ and $m_{h^0} > 67.5\,\text{GeV}$ with the region below 10 GeV being not excluded if $\tan\beta < 0.85$.

Chapter 5

Summary and Conclusion

An artificial neural network (ANN) has been used to search for neutral Higgs bosons in the context of both the Standard Model (SM) and a constrained Minimal Supersymmetric Model (MSSM). The search was performed in 54.1 pb^{-1} of data taken by the OPAL experiment at $\sqrt{s} = 183\,GeV$ in 1997. Both searches were performed in the *four jet* channel $e^+e^- \to HZ^0 \to b\bar{b}q\bar{q}$. Since the properties of the Standard Model Higgs boson and the lightest CP even neutral Higgs boson in the MSSM are nearly identical for the MSSM parameters considered in this thesis, the same analysis could be used for both the Standard Model and the MSSM. No statistically significant signal was observed above the expected background in either the SM or MSSM analyses. This analysis was the first application of an artificial neural network to an entire physics channel search in the OPAL Higgs group.

After an event quality selection and a loose precut to eliminate some of the two fermion and four fermion background signatures, ten variables were selected that gave good discrimination power between the signal and background channels. These variables were used as inputs to the ANN. The input variables made use of the four jet kinematic structure and the presence of b quark jets in the signal. The ANN used in the analysis was of the feed forward type and used backpropogation to update the weights of the network. The network consisted of three layers: an input layer with ten nodes, a single hidden layer with eleven nodes, and an output layer with three nodes. Each node of the input layer corresponded to one input variable. The

first output node was trained to select the signal events, the second output node to recognize the two fermion background, and the third output node to select the four fermion background. The network was trained using Monte Carlo events for 500 epochs and its performance was evaluated on a separate set of Monte Carlo events after each epoch. No evidence of overtraining was observed. The systematic errors inherent in the analysis were evaluated, and found to be 7.6% for the Higgs signal and 18.9% for the background.

A cut of 0.85 on the first output node of the ANN was used to select signal events in the OPAL data. The signal efficiency and background rejection of the network was evaluated using a separate *validation* set of Monte Carlo that had not been presented to the ANN before. The ANN performed well giving a peak efficiency of $(37.2 \pm 1.0(\text{stat.}) \pm 2.9(\text{syst.}))\%$ for a 85 GeV/c^2 Higgs boson. Seven candidate events were selected in the OPAL data which agrees with the $6.15 \pm 0.52(\text{stat.}) \pm 1.16(\text{syst.})$ events expected due to background processes alone. The Standard Model analysis was combined with the other decay channels of the Higgs boson at CERN in order to obtain a lower mass limit of 86.94 GeV/c^2 for a Standard Model Higgs boson at the 95% confidence level.

The ANN analysis was extended to search for the lightest neutral scalar Higgs boson of a constrained Minimal Supersymmetric Model through the production mechanism $e^+e^- \rightarrow h^0Z^0 \rightarrow b\bar{b}q\bar{q}$. Since the experimental signature of this channel is identical to the Standard Model four jet channel considered in this thesis, the same ANN could be used in the MSSM case. The results of this analysis were combined with the other MSSM search channels at CERN to produce exclusion plots in the $m_h - m_A$ plane. For $\tan\beta > 1$, the combined analysis excludes the CP-odd neutral MSSM Higgs boson at masses below 64.5 GeV. For $\tan\beta > 0.7$, however, the region $m_{A^0} < 13\,\text{GeV}$ is not excluded. Applying additional theoretical constraints to the model allows the 95% confidence limit exclusions to be extended to $m_{A^0} > 67.5\,\text{GeV}$ and $m_{h^0} > 67.5\,\text{GeV}$ with the region below 10 GeV being not excluded if $\tan\beta < 0.85$.

The future of the ANN based analysis in the OPAL Higgs group is bright. The Higgs search effort is entering an exciting region of the phase space as the OPAL centre of mass energy increases beyond 183 GeV. Fits to electroweak data from LEP, SLD, and the Tevatron [10] imply that the mass of the Standard Model Higgs boson is on the order of 100 GeV/c^2. Other theoretical analyses [57] imply that the mass of the lightest neutral CP even MSSM Higgs boson is less than about 130 GeV/c^2. OPAL will be able to search a considerable portion of this mass range (up to about $m_H = 100$ GeV/c^2), and so the prospects for discovery, or at least exclusion of the theoretically favoured Higgs mass range, are good. The evolution of the ANN analysis presented in this thesis will continue with the $\sqrt{s} = 192$ GeV and $\sqrt{s} = 200$ GeV runs at LEP.

Bibliography

[1] S. Glashow, "Partial symmetries of weak interactions," *Nucl. Phys.* **22** (1961) 579; S. Weinberg, "A model of leptons," *Phys. Rev. Lett.* **19** (1967) 1264; A. Salam, "Weak and electromagnetic interactions," in *Elementary Particle Theory*, W. Svartholm, ed., p. 367. Almquist and Wiksell, Stockholm, 1968.

[2] D.J. Gross and F. Wilczek *Phys. Rev.* **D8** (1973) 3633; D.J. Gross and F. Wilczek *Phys. Rev.* **D9** (1974) 908; F. Wilczek *Ann. Rev. Nucl. Part. Sci.* **32** (1982) 177.

[3] Cowsik *et. al. Phys. Rev. Lett.* **29** (1972) 669.

[4] The Particle Data Group, L. Montanet *et. al.*, *Eur. Phys. J.*, *Review of Particle Physics*, vol. 3. Springer-Verlag, Secaucus, NJ, 1998.

[5] P.W. Higgs *Phys. Rev.* **12** (1964) 132; F. Englert and R. Brout *Phys. Rev. Lett.* **13** (1964) 321; T.W.B. Kibble *Phys. Rev.* **D155** (1967) 1554.

[6] J.F. Gunion, H.E. Haber, G.L. Kane and S. Dawson, *The Higgs Hunters Guide.* Addison-Wesley Publishing Company, (Reading, MA), 1990.

[7] G. Altarelli, T. Sjöstrand and F. Zwirner, ed., *Physics at LEP2*, vol. 2. CERN, Feb., 1996.

[8] B. Ioffe and V. Khoze *Sov. J. Part. Nucl. Phys.* **9** (1978) 50.

[9] **OPAL** Collaboration, K. Ackerstaff *et. al. Eur. Phys. J.* **C1** (1998) 425.

[10] **LEP** Collaboration, D. Abbaneo *et. al.*, "A Combination of preliminary Electroweak measurements and constraints on the Standard Model,". CERN-EP-99-015.

[11] Y. Gol'fand and E. Likhtman *JETP Lett.* **13** (1971) 323; J. Wess and B. Zumino *Nucl. Phys.* **B70** (1974) 39; H. Haber and Y. Nir *Nucl. Phys.* **B335** (1990) 363.

[12] S. Dawson, "Introduction to the physics of Higgs bosons," **hep-ph/9411325**. .

[13] S. Weinberg *Phys. Rev.* **D13** (1976) 974; S. Weinberg *Phys. Rev.* **D19** (1979) 1277.

[14] S. P. Martin, "A Supersymmetry primer," **hep-ph/9709356**. .

[15] **OPAL** Collaboration, K. Ackerstaff *et. al.*, "A Search for neutral Higgs bosons in the MSSM and models with two scalar field doublets," *Eur. Phys. J.* **C5** (1998) 19, **hep-ex/9803019**. .

[16] P. Janot, "HZHA Higgs event generator. Version 2.00," in *Physics at LEP2* [7], p. 309.

[17] M. L. Mangano *et. al.*, "Event generators for discovery physics," hep-ph/9602203. .

[18] **UA1** Collaboration, G. Arnison *et. al. Phys. Lett.* **B122** (1983) 103.

[19] **OPAL** Collaboration, K. Ahmet *et. al. Nucl. Instrum. Methods* **A305** (1991) 275–319.

[20] **OPAL** Collaboration, G. Alexander *et. al. Phys. Lett.* **B374** (1996) 341.

[21] **OPAL** Collaboration, K. Akerstaff *et. al. Z. Phys.* **C73** (1997) 397.

[22] **OPAL** Collaboration, P. Allport *et. al. Nucl. Instrum. Methods* **A324** (1993) 34.

[23] **OPAL** Collaboration, P. Allport *et. al. Nucl. Instrum. Methods* **A346** (1994) 476.

[24] J. Carter *et. al. Nucl. Instrum. Methods* **A286** (1990) 99.

[25] O. Biebel *et. al. Nucl. Instrum. Methods* **A323** (1992) 169.

[26] H. Bethe, *Handbook der Physik 24/1.* J. Springer Verlag, Berlin, 1933.

[27] H. Mes *et. al. Nucl. Instrum. Methods* **A323** (1992) 169.

[28] G. Aguillion *et. al.*, "Thin scintillating tiles with high light yield for the OPAL endcaps," *Nucl. Instrum. Methods* **A417** (1998) 266–277.

[29] **OPAL** Collaboration, J. Hobbs *et. al. Nucl. Instrum. Methods* **A325** (1993) 494.

[30] **OPAL** Collaboration, S. Dado *et. al. Nucl. Instrum. Methods* **A252** (1986) 511.

[31] B. Denby, "The use of neural networks in high-energy physics,". FERMILAB-PUB-92-215-E.

[32] K. Gurney, *An Introduction to Neural Networks.* UCL Press, 1 Gunpowder Square, London, 1 ed., 1996.

[33] S. Amendolia, "Neural Networks," Yellow Report 94-06, CERN, 1994.

[34] B. Denby *et. al.* CDF Internal Report 1538, Fermilab, 1991.

[35] **CDF** Collaboration, B. Denby *et. al.*, "Investigation of a vlsi neural network chip as part of a secondary vertex trigger," *Nucl. Instrum. Methods* **A335** (1993) 296.

[36] T. Sjöstrand, "PYTHIA 5.7 and JETSET 7.4: Physics and Manual." LU-TP/95-20, CERN-TH.7112/93, Aug., 1995; T. Sjöstrand *Comp. Phys. Comm.* **39** (1986) 347; T. Sjöstrand and M. Bengtsson *Comp. Phys. Comm.* **43** (1987) 367.

[37] J. Fujimoto *et. al.*, "GRC4F v1.1: A four-fermion event generator for e^+e^- collisions," *Comp. Phys. Comm.* **100** (1997) 74; S. Kawabata *et. al.*, "GRC4F: User-Guide Version 2.1 Preliminary. GRACE manual: Automatic generation of tree amplitudes in Standard Models: Version 1.0." KEK-92-19, 1992.

[38] G. Marchesini *et. al. Comp. Phys. Comm.* **67** (1992) 465.

[39] F. Berends *et. al.*, "EXCALIBUR - a Monte Carlo Program to Evaluate all Four Fermion Processes at LEP200 and Beyond." INLO-PUB-12/94, July, 1994.

[40] J. Allison *et. al. Nucl. Instrum. Methods* **A317** (1992) 47.

[41] R. Bately, R. Hawkings, M. Hildreth, "`ODSIZL 2.11`, DST-Level Tracking Adjustments and Refit with Silicon Hits," OPAL Technical Report OPAL-TN/520, The OPAL Experiment, CERN, 1997.

[42] "Precision Luminosity for Z^0 Lineshape Measurements with the OPAL Silicon-Tungsten Luminometer," OPAL Technical Report OPAL-TN/528, The OPAL Experiment, CERN, 1997.

[43] S. Lloyd *et. al.*, "The OPAL Primer." An OPAL Experiment reference manual, July, 1997; C.P. Ward, D.R. Ward and C. Hartmann, "`WW` Manual." An OPAL Experiment program manual; P. Vikas *et. al.*, "`BT` code manual." An OPAL Experiment program manual.

[44] N. Brown and W. J. Stirling *Phys. Lett.* **B252** (1990) 657; S. Catani *et. al. Z. Phys.* **B** (1991) 432; D. S. S. Bethke, Z. Kunszt and W. J. Stirling *Nucl. Phys.* **B370** (1992).

[45] T.Omori, S.Asai, I.Nakamura and S.Yamashita, "A Matching Algorithm: MT package," OPAL Technical Report OPAL-TN/381, The OPAL Experiment, CERN, 1996.

[46] J. Bjorken and S. Brodsky *Phys. Rev.* **D1** (1970) 1416.

[47] **OPAL** Collaboration, G. Abbiendi *et. al.*, "Search for Higgs bosons in e^+e^- collisions at 183GeV," *Eur. Phys. J.* **C7** (1999) 407, `hep-ex/9811025`. .

[48] **OPAL** Collaboration, R. Akers *et. al. Z. Phys.* **C63** (1994) 197.

[49] **OPAL** Collaboration, R. Akers *et. al. Z. Phys.* **C66** (1995) 19; **OPAL** Collaboration, K. Ackerstaff *et. al. Z. Phys.* **C74** (1997) 1–17.

[50] R. Hawkings, "High performance b-tagging using 3d vertexing and neural networks," OPAL Technical Report OPAL-TN/515, The OPAL Experiment, CERN, Sept., 1997.

[51] "Search for neutral Higgs Bosons of the MSSM in e^+e^- collisions at $\sqrt{s}$ from 130 to 172 GeV," CERN Report CERN-PPE/97-071, The ALEPH Experiment, CERN, 1997.

[52] C. Peterson, T. Rögnvaldsson and L. Lönnblad, "`JETNET 3.0` - A Versatile Artificial Neural Network Package," Software Manual LT TP 93-29, CERN-TH.7135/94, Department of Theoretical Physics, University of Lund, Dec., 1993.

[53] **OPAL** Collaboration, G. Alexander *et. al. Z. Phys.* **C70** (1996) 357.

[54] **ALEPH** Collaboration, D. Buskulic *et. al. Phys. Lett.* **B357** (1995) 699; **ALEPH** Collaboration, D. Buskulic *et. al. Z. Phys.* **C62** (1994) 179; **DELPHI** Collaboration, P. Abreu *et. al. Z. Phys.* **C66** (1995) 323; **OPAL** Collaboration, R. Akers *et. al. Z. Phys.* **C60** (1993) 199.

[55] A.A. Faust OPAL Technical Report OPAL-TN/549, The OPAL Experiment, CERN, June, 1998.

[56] P. Bock, "Determination of exclusion limits for particle production using different decay channels with different efficiencies, mass resolutions, and backgrounds," Preprint HD-PY-96/05, Heidelberg University, Apr., 1996. To be submitted to Nucl. Instr. Meth..

[57] H. Haber and R. Hempfling *Phys. Rev. Lett.* **66** (1991) 1815; J. Ellis, G. Ridolfi and F. Zwirner *Phys. Lett.* **B257** (1991) 83; M. Berger *Phys. Rev.* **D41** (1990) 225.

Appendix A

The OPAL Collaboration

The OPAL Collaboration

G. Abbiendi[2], K. Ackerstaff[8], G. Alexander[23], J. Allison[16], N. Altekamp[5], K.J. Anderson[9], S. Anderson[12], S. Arcelli[17], S. Asai[24], S.F. Ashby[1], D. Axen[29], G. Azuelos[18,a], A.H. Ball[8], E. Barberio[8], R.J. Barlow[16], J.R. Batley[5], S. Baumann[3], J. Bechtluft[14], T. Behnke[27], K.W. Bell[20], G. Bella[23], A. Bellerive[9], S. Bentvelsen[8], S. Bethke[14], S. Betts[15], O. Biebel[14], A. Biguzzi[5], I.J. Bloodworth[1], P. Bock[11], J. Böhme[14], D. Bonacorsi[2], M. Boutemeur[33], S. Braibant[8], P. Bright-Thomas[1], L. Brigliadori[2], R.M. Brown[20], H.J. Burckhart[8], P. Capiluppi[2], R.K. Carnegie[6], A.A. Carter[13], J.R. Carter[5], C.Y. Chang[17], D.G. Charlton[1,b], D. Chrisman[4], C. Ciocca[2], P.E.L. Clarke[15], E. Clay[15], I. Cohen[23], J.E. Conboy[15], O.C. Cooke[8], J. Couchman[15], C. Couyoumtzelis[13], R.L. Coxe[9], M. Cuffiani[2], S. Dado[22], G.M. Dallavalle[2], R. Davis[30], S. De Jong[12], A. de Roeck[8], P. Dervan[15], K. Desch[27], B. Dienes[32,h], M.S. Dixit[7], J. Dubbert[33], E. Duchovni[26], G. Duckeck[33], I.P. Duerdoth[16], P.G. Estabrooks[6], E. Etzion[23], F. Fabbri[2], A. Fanfani[2], M. Fanti[2], A.A. Faust[30], L. Feld[10], F. Fiedler[27], M. Fierro[2], I. Fleck[10], A. Frey[8], A. Fürtjes[8], D.I. Futyan[16], P. Gagnon[7], J.W. Gary[4], J. Gascon[18], G. Gaycken[27], C. Geich-Gimbel[3], G. Giacomelli[2], P. Giacomelli[2], V. Gibson[5], W.R. Gibson[13], D.M. Gingrich[30,a], D. Glenzinski[9], J. Goldberg[22], W. Gorn[4], C. Grandi[2], K. Graham[28], E. Gross[26], J. Grunhaus[23], M. Gruwé[27], C. Hajdu[31] G.G. Hanson[12], M. Hansroul[8], M. Hapke[13], K. Harder[27], A. Harel[22], C.K. Hargrove[7], M. Harin-Dirac[4], M. Hauschild[8], C.M. Hawkes[1], R. Hawkings[27], R.J. Hemingway[6], G. Herten[10], R.D. Heuer[27], M.D. Hildreth[8], J.C. Hill[5], P.R. Hobson[25], A. Hocker[9], K. Hoffman[8], R.J. Homer[1], A.K. Honma[28,a], D. Horváth[31,c], K.R. Hossain[30], R. Howard[29], P. Hüntemeyer[27], P. Igo-Kemenes[11], D.C. Imrie[25], K. Ishii[24], F.R. Jacob[20], A. Jawahery[17], H. Jeremie[18], M. Jimack[1], C.R. Jones[5], P. Jovanovic[1], T.R. Junk[6], N. Kanaya[24], J. Kanzaki[24], D. Karlen[6], V. Kartvelishvili[16], K. Kawagoe[24], T. Kawamoto[24], P.I. Kayal[30], R.K. Keeler[28], R.G. Kellogg[17], B.W. Kennedy[20], D.H. Kim[19], A. Klier[26], T. Kobayashi[24], M. Kobel[3,d], T.P. Kokott[3], M. Kolrep[10], S. Komamiya[24], R.V. Kowalewski[28], T. Kress[4], P. Krieger[6], J. von Krogh[11], T. Kuhl[3], P. Kyberd[13], G.D. Lafferty[16], H. Landsman[22], D. Lanske[14], J. Lauber[15], I. Lawson[28], J.G. Layter[4], D. Lellouch[26], J. Letts[12], L. Levinson[26], R. Liebisch[11], B. List[8], C. Littlewood[5], A.W. Lloyd[1], S.L. Lloyd[13],

F.K. Loebinger[16], G.D. Long[28], M.J. Losty[7], J. Lu[29], J. Ludwig[10], D. Liu[12], A. Macchiolo[18], A. Macpherson[30], W. Mader[3], M. Mannelli[8], S. Marcellini[2], A.J. Martin[13], J.P. Martin[18], G. Martinez[17], T. Mashimo[24], P. Mättig[26], W.J. McDonald[30], J. McKenna[29], E.A. Mckigney[15], T.J. McMahon[1], R.A. McPherson[28], F. Meijers[8], P. Mendez-Lorenzo[33], F.S. Merritt[9], H. Mes[7], A. Michelini[2], S. Mihara[24], G. Mikenberg[26], D.J. Miller[15], W. Mohr[10], A. Montanari[2], T. Mori[24], K. Nagai[8], I. Nakamura[24], H.A. Neal[12,g], R. Nisius[8], S.W. O'Neale[1], F.G. Oakham[7], F. Odorici[2], H.O. Ogren[12], A. Okpara[11], M.J. Oreglia[9], S. Orito[24], G. Pásztor[31], J.R. Pater[16], G.N. Patrick[20], J. Patt[10], R. Perez-Ochoa[8], S. Petzold[27], P. Pfeifenschneider[14], J.E. Pilcher[9], J. Pinfold[30], D.E. Plane[8], P. Poffenberger[28], B. Poli[2], J. Polok[8], M. Przybycień[8,e], A. Quadt[8], C. Rembser[8], H. Rick[8], S. Robertson[28], S.A. Robins[22], N. Rodning[30], J.M. Roney[28], S. Rosati[3], K. Roscoe[16], A.M. Rossi[2], Y. Rozen[22], K. Runge[10], O. Runolfsson[8], D.R. Rust[12], K. Sachs[10], T. Saeki[24], O. Sahr[33], W.M. Sang[25], E.K.G. Sarkisyan[23], C. Sbarra[29], A.D. Schaile[33], O. Schaile[33], P. Scharff-Hansen[8], J. Schieck[11], S. Schmitt[11], A. Schöning[8], M. Schröder[8], M. Schumacher[3], C. Schwick[8], W.G. Scott[20], R. Seuster[14], T.G. Shears[8], B.C. Shen[4], C.H. Shepherd-Themistocleous[5], P. Sherwood[15], G.P. Siroli[2], A. Sittler[27], A. Skuja[17], A.M. Smith[8], G.A. Snow[17], R. Sobie[28], S. Söldner-Rembold[10,f], S. Spagnolo[20], M. Sproston[20], A. Stahl[3], K. Stephens[16], J. Steuerer[27], K. Stoll[10], D. Strom[19], R. Ströhmer[33], B. Surrow[8], S.D. Talbot[1], P. Taras[18], S. Tarem[22], R. Teuscher[9], M. Thiergen[10], J. Thomas[15], M.A. Thomson[8], E. Torrence[8], S. Towers[6], I. Trigger[18], Z. Trócsányi[32], E. Tsur[23], M.F. Turner-Watson[1], I. Ueda[24], R. Van Kooten[12], P. Vannerem[10], M. Verzocchi[8], H. Voss[3], F. Wäckerle[10], A. Wagner[27], C.P. Ward[5], D.R. Ward[5], P.M. Watkins[1], A.T. Watson[1], N.K. Watson[1], P.S. Wells[8], N. Wermes[3], D. Wetterling[11] J.S. White[6], G.W. Wilson[16], J.A. Wilson[1], T.R. Wyatt[16], S. Yamashita[24], V. Zacek[18], D. Zer-Zion[8]

[1]School of Physics and Astronomy, University of Birmingham, Birmingham B15 2TT, UK
[2]Dipartimento di Fisica dell' Università di Bologna and INFN, I-40126 Bologna, Italy
[3]Physikalisches Institut, Universität Bonn, D-53115 Bonn, Germany
[4]Department of Physics, University of California, Riverside CA 92521, USA
[5]Cavendish Laboratory, Cambridge CB3 0HE, UK
[6]Ottawa-Carleton Institute for Physics, Department of Physics, Carleton University, Ottawa, Ontario K1S 5B6, Canada
[7]Centre for Research in Particle Physics, Carleton University, Ottawa, Ontario K1S 5B6, Canada
[8]CERN, European Organisation for Particle Physics, CH-1211 Geneva 23, Switzerland
[9]Enrico Fermi Institute and Department of Physics, University of Chicago, Chicago IL 60637, USA
[10]Fakultät für Physik, Albert Ludwigs Universität, D-79104 Freiburg, Germany
[11]Physikalisches Institut, Universität Heidelberg, D-69120 Heidelberg, Germany
[12]Indiana University, Department of Physics, Swain Hall West 117, Bloomington IN 47405, USA
[13]Queen Mary and Westfield College, University of London, London E1 4NS, UK
[14]Technische Hochschule Aachen, III Physikalisches Institut, Sommerfeldstrasse 26-28, D-52056 Aachen, Germany

[15]University College London, London WC1E 6BT, UK
[16]Department of Physics, Schuster Laboratory, The University, Manchester M13 9PL, UK
[17]Department of Physics, University of Maryland, College Park, MD 20742, USA
[18]Laboratoire de Physique Nucléaire, Université de Montréal, Montréal, Quebec H3C 3J7, Canada
[19]University of Oregon, Department of Physics, Eugene OR 97403, USA
[20]CLRC Rutherford Appleton Laboratory, Chilton, Didcot, Oxfordshire OX11 0QX, UK
[22]Department of Physics, Technion-Israel Institute of Technology, Haifa 32000, Israel
[23]Department of Physics and Astronomy, Tel Aviv University, Tel Aviv 69978, Israel
[24]International Centre for Elementary Particle Physics and Department of Physics, University of Tokyo, Tokyo 113-0033, and Kobe University, Kobe 657-8501, Japan
[25]Institute of Physical and Environmental Sciences, Brunel University, Uxbridge, Middlesex UB8 3PH, UK
[26]Particle Physics Department, Weizmann Institute of Science, Rehovot 76100, Israel
[27]Universität Hamburg/DESY, II Institut für Experimental Physik, Notkestrasse 85, D-22607 Hamburg, Germany
[28]University of Victoria, Department of Physics, P O Box 3055, Victoria BC V8W 3P6, Canada
[29]University of British Columbia, Department of Physics, Vancouver BC V6T 1Z1, Canada
[30]University of Alberta, Department of Physics, Edmonton AB T6G 2J1, Canada
[31]Research Institute for Particle and Nuclear Physics, H-1525 Budapest, P O Box 49, Hungary
[32]Institute of Nuclear Research, H-4001 Debrecen, P O Box 51, Hungary
[33]Ludwigs-Maximilians-Universität München, Sektion Physik, Am Coulombwall 1, D-85748 Garching, Germany

[a] and at TRIUMF, Vancouver, Canada V6T 2A3
[b] and Royal Society University Research Fellow
[c] and Institute of Nuclear Research, Debrecen, Hungary
[d] on leave of absence from the University of Freiburg
[e] and University of Mining and Metallurgy, Cracow
[f] and Heisenberg Fellow
[g] now at Yale University, Dept of Physics, New Haven, USA
[h] and Depart of Experimental Physics, Lajos Kossuth University, Debrecen, Hungary.